

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
13 June 2002 (13.06.2002)

PCT

(10) International Publication Number
WO 02/46465 A2

(51) International Patent Classification⁷: **C12Q 1/68**

(21) International Application Number: PCT/GB01/05458

(22) International Filing Date:
10 December 2001 (10.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0030076.4 8 December 2000 (08.12.2000) GB
0103156.6 8 February 2001 (08.02.2001) GB
0125666.8 25 October 2001 (25.10.2001) GB

(71) Applicant (for all designated States except US): **OXFORD BIOMEDICA (UK) LIMITED** [GB/GB]; Medawar Centre, Robert Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **WHITE, Jonathan** [GB/GB]; Medawar Centre, Robert Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB). **MUNDY, Christopher, Robert** [AU/GB]; Medawar Centre, Robert Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB). **WARD, Neil, Raymond** [GB/GB]; Medawar Centre, Robert Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB). **KRIGE, David** [GB/GB]; Medawar Centre, Robert Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB). **KINGSMAN, Susan, Mary** [GB/GB]; Medawar Centre, Robert Robinson Avenue, The Oxford Science Park, Oxford OX4

4GA (GB). **HARRIS, Robert, Alan** [GB/GB]; Medawar Centre, Robert Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB). **RAYNER, William, Nigel** [GB/GB]; Medawar Centre, Robert Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB).

(74) Agents: **HALLYBONE, Huw, George** et al.; Carpmaels & Ransford, 43 Bloomsbury Square, London WC1A 2RA (GB).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ANALYSIS METHOD

(57) Abstract: This invention relates to novel methods for the identification of genes and gene products that are implicated in certain disease states. According to the invention, there is provided a method for the identification of a gene that is implicated in a specific disease or physiological condition, said method comprising the steps of comparing: i) the transcriptome or proteome of a first specialised cell type that is implicated in the disease or condition under first and second experimental conditions; with ii) the transcriptome or proteome of a second specialised cell type under said first and said second experimental conditions; and identifying as a gene implicated in the disease or physiological condition, a gene that is differentially regulated in the two specialised cell types under the first and second experimental conditions. The invention also relates to novel genes and gene products identified using these methods.

Analysis method

This invention relates to novel methods for the identification of genes and gene products that are implicated in certain disease states. The invention also relates to novel genes and gene products identified using these methods.

5 All publications, patents and patent applications cited herein are incorporated in full by reference.

One of the central goals in the field of gene expression is to understand and elucidate the relationship between a particular disease state and the gene expression pattern that defines and/or causes this disease state. Research has concentrated on differences in expression patterns between diseased and healthy tissues to elucidate the physiological mechanisms of disease. Identified differences in expression patterns
10 provide putative points for therapeutic intervention to reverse the disease phenotype. These differences also provide markers that are useful for diagnosis, and identify proteins for further investigation as agents implicated in the disease in question.

Conventional methods for the elucidation of mechanisms of disease tend to concentrate on the correlation of a disease state with altered levels of a particular protein. Such methods include techniques of
15 immunohistochemistry, the study of differential mRNA expression and the sequence analysis of particular proteins to identify mutations that are associated with a certain disease state.

Recently, research has concentrated on analysis of the transcriptomes of organisms and cell types that are considered to be of scientific interest. By "transcriptome" is meant the exact set of transcripts that are expressed in a cell. The emerging field of nucleic acid arrays is one field in which a large number of
20 powerful tools are being generated for the study of transcriptome variation between different tissue types. These tools are based on techniques originally pioneered by Schena *et al.*, 1995 (Science 270: 467-470) and Fodor *et al.*, 1991 (Science 251, 767-773) and facilitate the evaluation of variations in DNA or RNA sequences and of variations in expression levels from tissue samples and allow the identification and genotyping of mutations and polymorphisms in these sequences. The power of one such technique has
25 recently been demonstrated by Perou *et al.*, (Nature, 2000, 406:747-752), who generated molecular portraits of the transcriptomes of human breast tumours.

Over recent years, the so-called "genomics revolution" has allowed access to large portions of whole genomes, including the human genome. The amount of sequence information now available considerably facilitates the analysis of the results of experiments that aim to elucidate the differences between gene
30 expression in diseased and healthy tissues. As this information increases in scope and becomes more readily available, the study of the molecular mechanism of disease, and the elucidation of techniques for combatting these diseases will be considerably facilitated.

However, there are notable disadvantages associated with all methods that are currently employed for the

analysis of human disease. Many methods currently employed utilise established cell lines. Because these cells have been manipulated to allow their immortalisation in cell culture, the physiological situation in these cells is not considered by the present inventors to be generally representative of the authentic situation in equivalent cells *in vivo*. Furthermore, most of these methods tend to utilise a global strategy
5 for intervention, often ignoring the intricacies in gene expression that exist between different tissues. There thus remains a great need for the establishment of novel methods for the analysis of gene expression.

According to the invention, there is provided a method for the identification of a gene that is implicated in a specific disease or physiological condition, said method comprising the steps of:

- 10 a) comparing:
 - i) the transcriptome or proteome of a first specialised cell type that is implicated in the disease or condition under first and second experimental conditions; with
 - ii) the transcriptome or proteome of a second specialised cell type under said first and said second experimental conditions; and
- 15 b) identifying as a gene implicated in the disease or physiological condition, a gene that is differentially regulated in the two specialised cell types under the first and second experimental conditions.

Using this method, genes have been identified that respond to perturbations of cell physiology in a cell-specific rather than a generic fashion. The method of the invention exhibits significant advantages over
20 conventional methods of identifying genes that are implicated in disease.

Various groups have previously investigated mechanisms of physiological regulation, by comparing gene expression levels in the presence and absence of a physiological stimulus or challenge. Genes identified in a particular cell type as being expressed at different levels under different conditions are implicated as components of a pathway that is responsive to the altered conditions, or that is regulated differently under
25 the altered conditions. However, these methods exhibit a tendency to ignore patterns of gene expression that are physiologically relevant. This inclination is considered to result from a prejudice in the art that dictates that cells respond to changes in certain physiological conditions in a generic fashion, rather than in a cell specific fashion.

By "implicated in a specific disease or physiological condition" is meant that the gene has been found to
30 possess a distinct role in a pathway that is involved in susceptibility to, generation of or maintenance of a particular disease phenotype or physiological condition. As will be apparent to the skilled reader, any

point in any pathway may be the unique point at which a cell departs from the normal physiological response and generates a disease phenotype. Often the effect that is manifested as a disease is the result of a mutation event, in which a mutation occurs in the sequence of a gene encoding a protein that functions in a relevant physiological pathway.

- 5 There are numerous examples of diseases and conditions that may be studied using the method of the invention. Such pathological conditions include those that result from a change in the intrinsic nature of a cell (usually genetic) or from a change in the cellular microenvironment, either of which might be recapitulated in a laboratory setting. The methods may be applied to any disease or condition that is manifested in, or is generated in a specific cell type.
- 10 Examples of such conditions include changes in the cellular microenvironment, exposure to hormones, growth factors, cytokines, chemokines, inflammatory agents, toxins, metabolites, pH, pharmaceutical agents, hypoxia, anoxia, ischemia, imbalance of any plasma-borne nutrient [including glucose, amino acids, co-factors, mineral salts, proteins and lipids], osmotic stress, temperature [hypo and hyper-thermia], mechanical stress, irradiation [ionising or non-ionising], cell-extracellular matrix interactions,
- 15 cell-cell interactions, accumulations of foreign or pathological extracellular components, intracellular and extracellular pathogens [including bacteria, viruses, fungi and mycoplasma] and genetic perturbations [both epigenetic or mediated by mutation or polymorphism].

Examples of such diseases include cardiovascular disease, atherosclerosis, inflammatory conditions (including rheumatoid arthritis), cancer, ischemic disease, asthma, hematopoietic disorders, neurological

20 diseases including Parkinson's and Alzheimer's diseases, infectious disease and allergies.

One particular physiological response that has been used herein to illustrate the invention is the cellular response to hypoxia. The term "hypoxia" is intended to refer to an environment of reduced oxygen tension, as compared to the normal physiological environment for a particular organism, which is termed "normoxia". The prejudice in this technical field presents the view that there is a general, ubiquitous

25 response to hypoxia, mediated primarily at the level of mRNA (transcriptional initiation and post-transcriptional stabilisation).

In a variety of human diseases, cells are exposed to conditions of low oxygen tension, usually as a result of poor oxygen supply to the diseased area. For instance, tissue oxygenation plays a significant regulatory role in both apoptosis and in angiogenesis (Bouck *et al*, 1996, *Adv. Cancer Res.* 69:135-174; Bunn *et al*,

30 1996, *Physiol. Rev.* 76:839-885; Dor *et al*, 1997, *Trends Cardiovasc. Med.*, 7:289-294; Carmeliet *et al*, 1998, *Nature* 394:485-490). Apoptosis (see Duke *et al*, 1996, *Sci. American*, 80-87 for review) and growth arrest occur when cell growth and viability are reduced due to oxygen deprivation. Angiogenesis

(i.e. blood vessel growth, vascularization), is stimulated when hypooxygenated cells secrete factors that stimulate proliferation and migration of endothelial cells in an attempt to restore oxygen homeostasis (for review see Hanahan *et al*, 1996, *Cell*, 86:353-364).

Ischaemic disease pathologies involve a decrease in the blood supply to a bodily organ, tissue or body part generally caused by constriction or obstruction of the blood vessels. For example, solid tumours typically have a disorganised blood supply, leading to hypoxic regions. Other disease conditions involving hypoxia include stroke, atherosclerosis, retinopathy, acute renal failure, myocardial infarction, stroke and hair loss. Therefore, apoptosis and angiogenesis as induced by the ischaemic condition are also considered to be involved in these disease states. It is generally considered that understanding the mechanism by which cells respond to these diseases may be the key to the disease pathology and thus relevant to disease treatment.

In a different but related approach, it is now recognised that angiogenesis is necessary for tumour growth and that retardation of this process provide a useful tool in controlling malignancy and retinopathies. For example, neoangiogenesis is seen in many forms of retinopathy and in tumour growth. The ability to be able to induce tumourigenic cells to undergo apoptosis is an extremely desirable goal; particularly in the cancer field, it has been observed that apoptosis and angiogenesis-related genes provide potent therapeutic targets. It has also been observed that hypoxia plays a critical role in the selection of mutations that contribute to more severe tumourigenic phenotypes (Graeber *et al.*, 1996 *Nature*, 379(6560):88-91).

Early in the history of this field it was discovered that a transcription factor, HIF-1 α , is ubiquitously present in cells and is responsible for the induction of a number of genes in response to hypoxia. This protein is considered a master regulator of oxygen homeostasis (see, for example, Semenza, (1998) *Curr. Op. Genetics and Dev.* 8:588-594). Where HIF1 α is genetically knocked out, the hypoxia-inducible transcription of virtually all glycolytic enzymes has been shown to be inhibited. Glycolysis is an essential process which goes on in all mammalian cells. This finding is therefore consistent with previous work showing that when cells are exposed to conditions of hypoxia, they up-regulate glycolytic enzymes to enable ATP production, since oxidative phosphorylation is no longer feasible under conditions of low oxygen (Webster (1987) *Mol.Cell.Biochem*, 77: 19-28). Further support for a critical and general role of HIF1 α in the hypoxic response is demonstrated by the knockout mouse, which dies at day 10.5 of gestation. The same is true of the knockout of the ARNT protein, the dimerisation partner of HIF1 α .

For the first time, it is demonstrated herein that different tissues and cell types exhibit a very different response to hypoxia, at the level of the induction and repression of gene expression. This has allowed the

detailed elucidation of the mechanism of this particular physiological response, so paving the way for the development of improved therapeutic agents that target components of the response pathway in particular tissues. Although conventional approaches to the analysis of this mechanism have successfully identified numerous genes, because of the universal prejudice in the art that these components will all be induced/repressed similarly in all cell types, all the approaches suggested have hitherto been limited to the design of therapeutic agents that act in a global fashion.

The methods of the present invention therefore extend and add to previous work performed in this field, in that the discoveries made now allow the design of agents that target the hypoxic response in specific tissues. For example, it is known that brain and heart tissues die very rapidly after ischaemic insult. By using the method of the invention, it is quite possible that these tissues will be found to share common features in their response to hypoxia, that is different from other cell types. This might allow, for example, the design of a combination cardioprotective and neuroprotective agent effective against this subset of body tissues. Alternatively, the hypoxic response in these tissues might be found to be quite different. This information would then be taken into account when designing therapeutic countermeasures, in that an agent would be designed for the unique neurological or cardiological tissue concerned.

The method of the invention involves the comparison of the transcriptomes or proteomes of at least two specialised cell types under two different physiological or genetic conditions. By "transcriptome" is meant the exact set of transcripts that are expressed in a cell. The transcriptome thus has a qualitative element (the identity of individual gene transcripts) and a quantitative element (the proportion of each unique transcript in the total number of individual transcripts present in the cell at a particular moment). By "proteome" is meant the exact set of protein molecules that are expressed in a cell.

By "specialised cell type" is meant a cell type that has a restricted biochemical capacity and that can be unambiguously identified as possessing a unique set of biochemical and physiological functions. Preferably, the specialised cells are primary cells, and not cell lines or whole body tissues. Primary cells are cells that cannot proliferate indefinitely in culture. Primary cells can be derived from adult tissue, or from embryo tissue that is differentiated in culture to an adult cell or to a precursor of an adult cell that displays specialised characteristics.

Examples of preferred specialised cell types include cardiomyocytes, endothelial cells, sensory neurons, motor neurons, CNS neurons (all types), astrocytes, glial cells, schwann cells, mast cells, eosinophils, smooth muscle cells, skeletal muscle cells, pericytes, lymphocytes, tumor cells, monocytes, macrophages, foamy macrophages, granulocytes, synovial cells / synovial fibroblasts, epithelial cells (varieties from all

tissues/ organs). Examples of other suitable specialised cell types include vascular endothelial cells, smooth muscle cells (aortic, bronchial, coronary artery, pulmonary artery, etc), skeletal muscle cells, cardiomyocyte cells, fibroblasts (many types, such as synovial), keratinocytes, hepatocytes, dendritic cells, astrocytes, neurone cells (including mesencephalic, hippocampal, striatal, thalamic, hypothalamic, olfactory bulb, substantia nigra, locus coeruleus, cortex, dorsal root ganglia, superior cervical ganglia, sensory, motor, cerebellar cells), neutrophils, eosinophils, basophils, mast cells, monocytes, macrophage cells, erythrocytes, megakaryocytes, hematopoietic progenitor cells, hematopoietic pluripotent stem cells, any stem cells, any progenitor cells, epithelial cells, melanocytes, osteoblasts, osteoclasts, stromal cells, purkinje cells, T-cells, B-cells, synovial cells, pancreatic islet cells (alpha and beta), leukemia cells, lymphoma cells, tumour cells, retinal cells, adrenal chromaffin cells. As will be apparent to the skilled reader, it is not here possible to provide an exhaustive list of specialised cell types that may be studied according to the methods of the present invention.

Intended as being included within the method of the invention is the possibility of using, as two different specialised cell types, two different physiological states of the same cell type, for example, activated and resting macrophages.

The transcriptomes of the specialised cell types are compared under different experimental conditions. The term "experimental conditions" is used broadly in this context and is intended to embrace any physiological or genetic conditions to which a cell type may be exposed. The intention of the method is to compare the transcriptomes or proteomes of the cell types under different experimental conditions that have a physiological relevance. Accordingly, the state of the transcriptome or proteome under one set of experimental conditions will generally act as a control against which the transcriptome or proteome may be compared under a second set of experimental conditions. Any distinct physiologically-relevant conditions may therefore be of interest.

Examples of suitable physiological experimental conditions include conditions under which the cell is submitted to a physiological, mechanical, temperature, chemical, toxic or pharmaceutical stress. One example is hypoxia, defined herein as a physiological state in which oxygen demand by the cell exceeds its supply to the cell. The transcriptome or proteome under this set of experimental conditions may be compared to the transcriptome or proteome under conditions of normoxia, when oxygen supply is in concordance with the demand by the cell.

The transcriptomes or proteomes may also be compared under different genetic conditions. By "genetic conditions" is meant that the genotype of the compared cell populations contains a different genetic component. This may be the presence of one or more different, non-endogenous nucleic acid molecules in

the cell, herein referred to collectively as "genetic elements". Such genetic element(s) may potentially be incorporated into the genome of the cell, or alternatively may exist as a separate genetic entity, for example, as an extra-chromosomal element such as a plasmid or episome. Alternatively, the genome may have been perturbed by external intervention, for example, to increase or decrease the expression of a particular gene or genes. A number of variations on this theme are possible, including the overexpression of a genetic element via the administration of the functional gene, the overexpression of a genetic element via the administration of a regulator of the functional gene (such as, for example, a transcription factor [either natural or artificially constructed via the fusion of a DNA binding domain with an activator domain]), the inhibition of the expression of a functional gene (for example, using antisense RNA or ribozymes), the inhibition of the expression of a functional gene (for example, using a transdominant protein) and the inhibition of the expression of a functional gene (for example, using a repressor protein that is either natural or artificially constructed from a DNA binding protein fused to a repressor domain).

A particular example of a genetic perturbation as envisaged herein, that forms one preferred embodiment of the method of the present invention, is the so-called "Smartomics" technology that forms the basis for co-pending, co-owned International patent application PCT/GB01/00758. According to this technology, a heterologous nucleic acid is introduced into a primary cell to augment a specific natural physiological response. "Smartomics" may be applied to the current invention by measuring and comparing cellular responses to a heterologous gene in two or more distinct cell types, both with and without the natural physiological stimulus. Lentivirus technology is used to introduce the heterologous nucleic acid molecule in such a way that there is negligible perturbation of endogenous gene expression. For this reason, this technology exhibits significant benefits over conventional technology of a similar nature, since the prior art methods are generally invasive, having downstream effects other than the simple introduction of the heterologous nucleic acid molecule. The Smartomics technology allows much more precise measurements to be taken of the effect of introducing the heterologous nucleic acid.

The method of the invention allows the identification of genes that are implicated in a specific disease or physiological condition. The genes identified in this way are candidate targets for antagonists or agonists that modulate disease states pertinent to that specialised cell type. This allows the development of selective agonists and antagonists, rather than broad spectrum agonists and antagonists. This approach thus adds value in the selective treatment of disease. Furthermore the identified genes are associated with regulatory elements that provide alternative and additional candidate targets for exploitation for the delivery of gene products to that cell in a cell-specific fashion. The genes and regulatory elements identified according to the method of the invention can be used directly in therapeutic applications via gene therapy, via recombinant protein methods or via chemical mimetics or as targets for the

development of agonists and antagonists such as antibodies, small chemical molecules, peptides, regulatory nucleic acids.

The step of comparison of the transcriptomes or proteomes of the first and second specialised cell types under first and second experimental conditions may be effected using any approach that allows the
5 quantitative comparison of gene or protein expression, and a number of such means will be known to those of skill in the art. Such experiments have only become possible in recent years, due to certain advances in technology that have allowed the large scale, high throughput analysis of gene expression.

One example of a method that allows the comparison of the transcriptome of a specific cell type with a second or subsequent transcriptome involves the generation of a set of clones that represent all the
10 transcripts expressed in a cell under the conditions in which the cell is maintained. This may be done by constructing a cDNA library, in which copies of all mRNA transcripts expressed in the cell are cloned into a suitable vector for subsequent analysis.

Such libraries may be normalised cDNA libraries, in which the distribution of genes in the library has been biased to reduce the number of clones that represent genes with large numbers of transcripts (such
15 as, for example, beta-actin) and thus reduce the repetitive nature of the library. Normalisation thus acts to reduce the frequency of genes expressed at high levels and to enhance the frequency of genes expressed at low levels (see de Fatima Bonaldo *et al.*, Genome Research 6: 791-806 (1996)).

Libraries may also be subtracted cDNA libraries, in which the distribution of genes is manipulated to remove genes that are expressed in both mRNA populations used to construct the library. The
20 commercially-available PCR Select kit (Clontech, Inc) is an example of a system useful to generate such libraries.

cDNA clones generated as reflective of the transcriptome of a specific cell type may then be amplified, and processed to evaluate the identity of the nucleic acid clones. For example, multiple clones may be picked and used as template for PCR amplification. The PCR products may then be arrayed onto
25 membranes or glass slides to create nucleic acid arrays. For expression profiling, these arrays are then hybridised to complex nucleic acid probes in order to quantitate the abundance of individual genes contained in the probes.

A recent summary of nucleic acid array technology that is useful in the analysis of the transcriptome of a cell population is provided in Nature Genetics, (1999) (21 suppl; 1-61). There are various types of array
30 technology currently used, including "microarrays", or "chips", which are high density cDNA arrays produced on glass slides, often produced using photolithography. A second type of array is the

"macroarray", which is an array with sub-millimetre spot-spot distances produced on a nylon membrane. One example of this type of array are the nylon-based microarrays sold commercially by Research Genetics Inc. (termed Research Genetics Human GeneFilters) that each contain 5,300 cDNA fragments of known identity. The whole series of arrays covers some 35,000 cDNA fragments. This particular array
5 system (and others like it) allow the identification of transcripts that are down-regulated, as well as those that are up-regulated, since the range of genes used to manufacture the arrays are not biased.

The step of comparison may be effected by utilising subtracted cDNA libraries. Using this approach, the transcriptome of one specialised cell type under first experimental conditions is subtracted against the transcriptome under second experimental conditions. This reveals the differences in expression under the
10 two experimental conditions tested. When this is performed for both specialised cell types, the differential regulation of gene expression under the two experimental conditions is revealed.

The step of comparison is through the detection of genes that are differentially regulated in the two specialised cell types examined under the first and second experimental conditions. As an example, a human cardiomyoblast (cell type A) and a human macrophage (cell type B) may be placed at the same
15 temperature and at a high oxygen tension (first experimental conditions [1]). Cells from the same cell types are also incubated at this temperature, yet under conditions of low oxygen tension (second experimental conditions [2]). In this simple example, there are then a minimum of four combinations of cell type and condition, A[1], B[1], A[2] and B[2]. "Snapshots" are taken of the transcriptomes of both cell types under the "normoxic" and the "hypoxic" experimental conditions, by preparing messenger
20 RNA from all four combinations. Differences in the regulation of genes can then be analysed, for example, using a process of subtractive hybridisation.

The mechanism of transcriptome comparison in the above example may be as follows. Subtracted cDNA libraries are separately prepared for hypoxic macrophages and cardiomyoblasts; for both cell types, their cDNA under normoxic conditions is subtracted against their cDNA under hypoxic conditions. This might
25 be effected by harvesting RNA from cells both in normoxia and hypoxia, and preparing cDNA. Subtractive hybridization, optionally including suppression PCR, may then be performed to remove genes from the hypoxic cell cDNA which are also present in cDNA from normoxic cells. Insert DNA from these subtracted libraries can then be amplified and arrayed onto duplicate membranes. Quantitative hybridization with pre-library cDNA material (normoxia and hypoxia) then allows the comparison of
30 differentially-expressed clones in the two cell types. The clones representing hypoxia-inducible genes may be then be identified, for example, by sequencing.

Other techniques that are suitable for the analysis of the transcriptome of a specific cell type include serial analysis of gene expression (SAGE; Velculescu *et al.*, Science (1995) 270: 484-487), Selective amplification via biotin- and restriction-mediated enrichment (SABRE) (Lavery *et al.*, (1997), *PNAS USA* 94: p6831-6836); Differential display (for example, indexing differential display reverse transcriptase
5 polymerase chain reaction (DDRT-PCR; Mahadeva *et al.* (1998) *J. Mol.Biol.* 284, 1391-1398)); representational difference analysis (RDA) (Hubank (1999) *Methods in Enzymology* 303: 325-349); differential screening of cDNA libraries (see Sagerstrom *et al.* (1997) *Annu. Rev. Biochem.* 66: 751-783); "Advanced Molecular Biology", R.M. Twyman (1998) Bios Scientific Publishers, Oxford; "Nucleic Acid Hybridization", M. L. M. Anderson (1999) Bios Scientific Publishers, Oxford); Northern blotting; RNase
10 protection assays; S1-nuclease protection assays; RT-PCR; real time RT-PCR (Taq-man); EST sequencing; massively parallel signature sequencing (MPSS); and sequencing by hybridisation (SBH) (see Drmanac R. *et al* (1999), *Methods in Enzymology* 303:165-178). Many of these techniques are reviewed in "Comparative gene-expression analysis" *Trends Biotechnol.* 1999 Feb;17(2):73-8.

Methods such as these have been applied widely to study mechanisms of biological response. In
15 particular, microarrays have been used widely to compare gene expression levels between normal and diseased tissue. More typically, however, comparisons are performed to detect changes in gene expression that are associated with specific aspects of disease progression or pathology. For instance, a study of prostate cancer would examine changes associated with the step-wise progression to full malignancy or the dependence on androgens for growth.

20 Transcriptome analysis is complemented by the analysis of the complete protein make-up of a cell, referred to as proteomics. The use of two dimensional SDS-PAGE gels in combination with amino acid sequencing by mass spectrometry is currently the most widely-used technique in this field (see "Proteomics to study genes and genomes" Akhilesh Pandey and Matthias Mann, (2000), *Nature* 405: 837-846). Additionally, the recent developments in the field of protein and antibody arrays now allow the
25 simultaneous detection of a large number of proteins. For example, low-density protein arrays on filter membranes, such as the universal protein array system (Ge H, (2000) *Nucleic Acids Res.* 28(2), e3) allow imaging of arrayed antigens using standard ELISA techniques and a scanning charge-coupled device (CCD) detector. Immuno-sensor arrays have also been developed that enable the simultaneous detection of clinical analytes. It is now possible using protein arrays, to profile protein expression in bodily fluids,
30 such as in sera of healthy or diseased subjects, as well as in patients pre- and post-drug treatment.

Antibody arrays also facilitate the extensive parallel analysis of numerous proteins that are hypothetically implicated in a disease or particular physiological state. A number of methods for the preparation of antibody arrays have recently been reported (see Cahill, *Trends in Biotechnology*, 2000 7:47-51).

It is not the intention here to review studies that have been conducted in this area previously. However, one example of a physiological condition that has already received considerable attention is the response to hypoxia. Several patent applications have now been published that involve an examination of the genetic response to hypoxia (see WO00/12139, Quark Biotech, Inc.; WO00/12525, Quark Biotech, Inc.;
5 WO99/09049, Quark Biotech, Inc.; WO99/09046, Quark Biotech, Inc.; WO99/48916, The Board of Trustees of the Leland Stanford Jr. University). These patent applications generally utilise methods of subtractive hybridisation and differential expression gene microarray analysis to examine this genetic response in certain cell lines. The studies have implicated specific genes as being either repressed or induced under hypoxic conditions as compared to their expression under normoxic conditions. These
10 genes are taught as being useful generally in all cell types, being involved in the (generic) hypoxic response.

Significantly, the present invention extends this work, and, indeed, defines a significant advance over similar work that has been performed on the genetic mechanisms that act in response to other physiological or genetic stimuli. The present inventors, using the novel methods disclosed herein, have
15 discovered that far from being generic, the cellular response to many physiological conditions differs markedly between different cell types. The cellular response that has been studied in order to illustrate this finding is the response to hypoxia. From these results, it has been inferred herein, quite reasonably, that far from being generic, cellular response mechanisms differ widely, depending on cell type.

This discovery has far-reaching implications as regards the design of therapeutic agents that are effective
20 to counter a disease or physiological condition. For example, an agent that is effective to prevent the drastic effects of hypoxia in a neurone (the effects of which include stroke) might be totally ineffective in countering the same effects in a cardiomyocyte (chronic ischemic heart disease). Through analysing the mechanism of the hypoxic response in different cell types, it may be, in contrast to the example given above, that a particular gene is involved in the hypoxic response in both cardiomyocytes and neurones.
25 Were this to be the case, this would allow the design of a combined medicament, for example, a combined cardioprotective and neuroprotective agent. There thus remains a great need for the identification of proteins implicated in the physiological mechanism of hypoxia.

According to a further aspect of the invention, there are provided genes and proteins that are identified using a method according to any one of the above-described aspects of the invention. Certain proteins,
30 whose sequences are identified herein as SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 63, 67, 69, 73, 75, 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129, 131, 133, 137, 139, 141, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 and 209, are functionally annotated for the first time. At present, all of these sequences are only

identified as "hypothetical proteins" in the public databases. Each and every one of these sequences forms an embodiment of this aspect of the invention.

The invention also includes proteins whose amino acid sequences are encoded by a nucleic acid sequence recited in various cDNAs and ESTs deposited in the public databases, or encoded by a gene identified
 5 from such an EST. These cDNAs and ESTs are presented herein as SEQ ID Nos: 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216. At present, all of these cDNA and EST sequences are functionally
 10 unannotated in the public databases. Each and every one of these sequences forms an embodiment of this aspect of the invention.

One embodiment of this aspect of the invention provides substantially purified polypeptide, which polypeptide:

i) comprises the amino acid sequence as recited in any one of SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13,
 15 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 63, 67, 69, 73, 75, 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129, 131, 133, 137, 139, 141, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 or 209;

ii) has an amino acid sequence encoded by a nucleic acid sequence recited in any one of SEQ
 20 ID Nos: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216 or encoded by a gene identified from an EST recited in any one of these SEQ
 25 ID Nos;

iii) is a fragment of a polypeptide according to i) or ii), provided that said fragment retains a biological activity possessed by the full length polypeptide of i) or ii), or has an antigenic determinant in common with the polypeptide of i) or ii); or

iv) is a functional equivalent of a polypeptide of i), ii) or (iii).

30 The polypeptide sequences recited in SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 63, 67, 69, 73, 75, 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129, 131, 133, 137, 139, 141, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 and 209 were, prior to the present disclosure, unannotated in the literature and public sequence

databases. Accordingly, until now, no biological function has been attributed to these polypeptide sequences; each of these sequences is generally labelled in the databases as a "hypothetical protein". The methods of the present invention, described above, have now elucidated a biological function for these polypeptides, in that they have been found to be differentially regulated under physiological conditions of
5 hypoxia.

These discoveries allow the development of regulators, such as small drug molecules, that affect the activity of these polypeptides, so allowing diseases and physiological conditions that are caused by hypoxia, or in which hypoxia has been implicated, to be treated. These discoveries also allow the development of diagnostic agents that are suitable for the detection of hypoxia in biological tissues and,
10 through the identification of mutations and polymorphisms (such as SNPs) within genes coding for the proteins implicated herein, allows the assessment of an individual's risk of being susceptible to diseases and physiological conditions in which hypoxia is implicated.

The biological activity of polypeptides whose sequences are listed in SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 63, 67, 69, 73, 75,
15 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129, 131, 133, 137, 139, 141, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 and 209 has been found to be hypoxia-regulated. The expression of some of these polypeptides has been found to be induced under conditions of hypoxia, whilst the expression of other polypeptides has been found to be repressed. By "hypoxia-induced" is meant that the polypeptide is expressed at a higher level when a cell is exposed to hypoxic conditions as compared to its
20 expression level under normoxic conditions. By "hypoxia-repressed" is meant that the polypeptide is expressed at a lower level when a cell is exposed to hypoxic conditions as compared to its expression level under normoxic conditions.

The following polypeptides have been found to be hypoxia-induced: those polypeptides whose amino acid sequence is recited in SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 63, 67,
25 69, 73, 75, 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129, 131, 133, 137, 139 and 141; and those polypeptides whose amino acid sequence is encoded by a nucleic acid sequence recited in SEQ ID Nos.: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142 and 144 or is encoded by a gene identified from an EST recited in any
30 one of these SEQ ID Nos..

The following polypeptides have been found to be hypoxia-repressed: those polypeptides whose amino acid sequence is recited in SEQ ID Nos.: 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 and 209; and those polypeptides whose amino acid sequence

is encoded by a nucleic acid sequence recited in SEQ ID Nos.: 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216, or encoded by a gene identified from an EST recited in any one of these SEQ ID Nos.

- 5 For the purposes of this document, the term "hypoxia" should be taken to mean an environment of oxygen tension such that the oxygen content is between about 5% and 0.1% (v/v). In most cases, hypoxic tissue will have an oxygen content that is less than or equal to about 2%. The term "normoxia" should be taken to mean conditions comprising a normal level of oxygen for the environment concerned. Normoxic tissue typically has an oxygen content above about 5%.
- 10 The polypeptide sequences whose amino acid sequence is encoded by a nucleic acid sequence recited in SEQ ID Nos: 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216, or whose amino acid sequence
- 15 is encoded by a gene identified from an EST recited in any one of these SEQ ID Nos., were also, prior to the present disclosure, unannotated in the literature and public sequence databases, meaning that until now, no biological function has been attributed to these polypeptide sequences.

- The sequences in this group fall into a number of different categories. The first of these are cDNA clones, for which a protein sequence has not been predicted by the depositor. A second category is expressed
- 20 sequence tag (EST) sequences that are represented in the UniGene database (<http://www.ncbi.nlm.nih.gov/UniGene/>), which contain modest or weak homology to known proteins when translated. ESTs are single-pass sequence files of the 5' region of an organism's expressed genome as accessed via a force cloned cDNA library. EST sequences tend to be short and as a general rule are error-prone. UniGene (see <http://www.ncbi.nlm.nih.gov/Web/Newsltr/aug96.html> for review) is an
- 25 experimental system for automatically partitioning these EST sequences into a non-redundant set of gene-oriented clusters. Each UniGene cluster contains sequences that represent a unique gene, as well as related information such as the tissue types in which the gene has been expressed and map location. A third category of hits identified by the methods described herein is EST sequences that are contained in
- 30 public databases. The fourth and final category encompasses singleton EST sequence entries that are not incorporated as entries in the Unigene database and that only appear as single entries in the public databases.

The methods of the present invention, described above, have now elucidated a biological function for polypeptides that are encoded by genes incorporating cDNA and EST sequences that fall into the four categories set out above, in that these sequences have been found to be differentially regulated under physiological conditions of hypoxia. Such polypeptides may have an amino acid sequence that is encoded
 5 by a nucleic acid sequence recited in any one of SEQ ID Nos: 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216. However, the EST sequences in particular may not be part of the actual coding
 10 sequence for a gene, often representing regulatory regions of the gene, or regions that are transcribed, but not translated into polypeptide. Accordingly, this aspect of the invention also includes polypeptides that are encoded by a gene identified from an EST recited in any one of SEQ ID Nos: 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216;

Polypeptides of this aspect of the invention are intended to include fragments of polypeptides according to i) or ii) as defined above, provided that the fragment retains a biological activity that is possessed by the full length polypeptide of i) or ii), or has an antigenic determinant in common with the polypeptide of
 20 i) or ii). As used herein, the term "fragment" refers to a polypeptide having an amino acid sequence that is the same as part, but not all, of an amino acid sequence as recited in any one of SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 63, 67, 69, 73, 75, 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129, 131, 133, 137, 139, 141, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 and 209, an amino acid sequence that is encoded by a
 25 nucleic acid sequence recited in any one of SEQ ID Nos. 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216, or an amino acid sequence that is encoded by a gene that is linked to a nucleic acid sequence
 30 recited in any one of these SEQ ID Nos. The fragments should comprise at least n consecutive amino acids from the sequence and, depending on the particular sequence, n preferably is 7 or more (for example, 8, 10, 12, 14, 16, 18, 20 or more). Small fragments may form an antigenic determinant.

Such fragments may be isolated fragments, that are not part of or fused to other amino acids or polypeptides, or they may be comprised within a larger polypeptide, of which they form a part or region.

When comprised within a larger polypeptide, a fragment of the invention most preferably forms a single continuous region. For instance, certain preferred embodiments relate to a fragment having a pre- and/or pro- polypeptide region fused to the amino terminus of the fragment and/or an additional region fused to the carboxyl terminus of the fragment. However, several fragments may be comprised within a single
5 larger polypeptide.

The polypeptides of the present invention or their immunogenic fragments (comprising at least one antigenic determinant) can be used to generate ligands, such as polyclonal or monoclonal antibodies, that are immunospecific for the polypeptides. Such antibodies may be employed to isolate or to identify clones that express a polypeptide according to the invention or, for example, to purify the polypeptide by
10 affinity chromatography. Such antibodies may also be employed as diagnostic or therapeutic aids, amongst other applications, as will be apparent to the skilled reader.

The term "immunospecific" means that an antibody has substantially greater affinity for a polypeptide according to the invention than their affinity for related polypeptides. As used herein, the term "antibody" is intended to include intact molecules as well as fragments thereof, such as Fab, F(ab')₂ and scFv, which
15 are capable of binding to the antigenic determinant in question.

The invention also includes functional equivalents of a polypeptide of i), ii) or (iii) as recited above. A functionally-equivalent polypeptide according to this aspect of the invention may be a polypeptides that is homologous to a polypeptide whose sequence is explicitly recited herein. Two polypeptides are said to be "homologous" if the sequence of one of the polypeptides has a high enough degree of identity or
20 similarity to the sequence of the other polypeptide for the skilled person to determine that they are similar in origin and function. Preferably, homology is used to refer to sequence identity. "Identity" indicates that at any particular position in the aligned sequences, the amino acid residue is identical between the sequences. "Similarity" indicates that, at any particular position in the aligned sequences, the amino acid residue is of a similar type between the sequences. Degrees of identity and similarity can be readily
25 calculated according to methods known in the art (see, for example, Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing. Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993). Advantageously, the BLAST algorithm is employed, with parameters set to default values. The BLAST algorithm is described in detail at http://www.ncbi.nih.gov/BLAST/blast_help.html, which is incorporated herein by reference. The search
30 parameters are defined as follows, and are advantageously set to the defined default parameters.

Advantageously, "substantial homology" when assessed by BLAST equates to sequences which match with an EXPECT value of at least about 7, preferably at least about 9 and most preferably 10 or more. The default threshold for EXPECT in BLAST searching is usually 10.

BLAST (Basic Local Alignment Search Tool) is the heuristic search algorithm employed by the programs blastp, blastn, blastx, tblastn, and tblastx; these programs ascribe significance to their findings using the statistical methods of Karlin and Altschul (see http://www.ncbi.nih.gov/BLAST/blast_help.html) with a few enhancements. The BLAST programs were tailored for sequence similarity searching, for example to
5 identify homologues to a query sequence. The programs are not generally useful for motif-style searching. For a discussion of basic issues in similarity searching of sequence databases, see Altschul et al. (1994) Nature Genetics 6:119-129.

The five BLAST programs available at <http://www.ncbi.nlm.nih.gov> perform the following tasks:

blastp compares an amino acid query sequence against a protein sequence database;

10 blastn compares a nucleotide query sequence against a nucleotide sequence database;

blastx compares the six-frame conceptual translation products of a nucleotide query sequence (both strands) against a protein sequence database;

tblastn compares a protein query sequence against a nucleotide sequence database dynamically translated in all six reading frames (both strands).

15 tblastx compares the six-frame translations of a nucleotide query sequence against the six-frame translations of a nucleotide sequence database.

BLAST uses the following search parameters:

HISTOGRAM Display a histogram of scores for each search; default is yes. (See parameter H in the BLAST Manual).

20 DESCRIPTIONS Restricts the number of short descriptions of matching sequences reported to the number specified; default limit is 100 descriptions. (See parameter V in the manual page). See also EXPECT and CUTOFF.

ALIGNMENTS Restricts database sequences to the number specified for which high-scoring segment pairs (HSPs) are reported; the default limit is 50. If more database sequences than this happen to satisfy
25 the statistical significance threshold for reporting (see EXPECT and CUTOFF below), only the matches ascribed the greatest statistical significance are reported. (See parameter B in the BLAST Manual).

EXPECT The statistical significance threshold for reporting matches against database sequences; the default value is 10, such that 10 matches are expected to be found merely by chance, according to the stochastic model of Karlin and Altschul (1990). If the statistical significance ascribed to a match is greater
30 than the EXPECT threshold, the match will not be reported. Lower EXPECT thresholds are more

stringent, leading to fewer chance matches being reported. Fractional values are acceptable. (See parameter E in the BLAST Manual).

CUTOFF Cutoff score for reporting high-scoring segment pairs. The default value is calculated from the EXPECT value (see above). HSPs are reported for a database sequence only if the statistical significance
5 ascribed to them is at least as high as would be ascribed to a lone HSP having a score equal to the CUTOFF value. Higher CUTOFF values are more stringent, leading to fewer chance matches being reported. (See parameter S in the BLAST Manual). Typically, significance thresholds can be more intuitively managed using EXPECT.

MATRIX Specify an alternate scoring matrix for BLASTP, BLASTX, TBLASTN and TBLASTX. The
10 default matrix is BLOSUM62 (Henikoff & Henikoff, 1992). The valid alternative choices include: PAM40, PAM120, PAM250 and IDENTITY. No alternate scoring matrices are available for BLASTN; specifying the MATRIX directive in BLASTN requests returns an error response.

STRAND Restrict a TBLASTN search to just the top or bottom strand of the database sequences; or restrict a BLASTN, BLASTX or TBLASTX search to just reading frames on the top or bottom strand of
15 the query sequence.

FILTER Mask off segments of the query sequence that have low compositional complexity, as determined by the SEG program of Wootton & Federhen (1993) Computers and Chemistry 17:149-163, or segments consisting of short-periodicity internal repeats, as determined by the XNU program of Claverie & States (1993) Computers and Chemistry 17:191-201, or, for BLASTN, by the DUST program
20 of Tatusov and Lipman (see <http://www.ncbi.nlm.nih.gov>). Filtering can eliminate statistically significant but biologically uninteresting reports from the blast output (e.g., hits against common acidic-, basic- or proline-rich regions), leaving the more biologically interesting regions of the query sequence available for specific matching against database sequences.

Low complexity sequence found by a filter program is substituted using the letter "N" in nucleotide
25 sequence (e.g., "NNNNNNNNNNNNNN") and the letter "X" in protein sequences (e.g., "XXXXXXXXXX").

Filtering is only applied to the query sequence (or its translation products), not to database sequences. Default filtering is DUST for BLASTN, SEG for other programs.

It is not unusual for nothing at all to be masked by SEG, XNU, or both, when applied to sequences in
30 SWISS-PROT, so filtering should not be expected to always yield an effect. Furthermore, in some cases, sequences are masked in their entirety, indicating that the statistical significance of any matches reported against the unfiltered query sequence should be suspect.

NCBI-gi Causes NCBI gi identifiers to be shown in the output, in addition to the accession and/or locus name.

Most preferably, sequence comparisons are conducted using the simple BLAST search algorithm provided at <http://www.ncbi.nlm.nih.gov/BLAST>.

- 5 Alternatively, sequence homology may be determined by algorithms such as FastA, available at <http://biology.ncsa.uiuc.edu/BW30/BW.cgi>. FastA is considered to be superior to BLAST for alignment of short sequences. Advantageously, the FastA algorithm is employed using default parameters at <http://biology.ncsa.uiuc.edu/BW30/BW.cgi>.

- Typically, greater than 50% identity between two polypeptides is considered to be an indication of
10 functional equivalence, provided that either the biological activity of the polypeptide is retained or the polypeptides possess an antigenic determinant in common. Preferably, a functionally equivalent polypeptide according to this aspect of the invention exhibits a degree of sequence identity with a polypeptide sequence explicitly identified herein, or with a fragment thereof, of greater than 50%. More preferred polypeptides have degrees of identity of greater than 60%, 70%, 80%, 90%, 95%, 98% or 99%,
15 respectively.

- Functionally-equivalent polypeptides according to the invention are therefore intended to include natural biological variants (for example, allelic variants or geographical variations within the species from which the polypeptides are derived) and mutants (such as mutants containing amino acid substitutions, insertions or deletions) of the polypeptides whose sequences are explicitly recited herein. Such mutants may include
20 polypeptides in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; among the basic residues Lys and Arg; or among the aromatic residues Phe and Tyr.

- 25 Particularly preferred are variants in which several, i.e. between 5 and 10, 1 and 5, 1 and 3, 1 and 2 or just 1 amino acids are substituted, deleted or added in any combination. Especially preferred are silent substitutions, additions and deletions, which do not alter the properties and activities of the protein. Also especially preferred in this regard are conservative substitutions. "Mutant" polypeptides also include polypeptides in which one or more of the amino acid residues include a substituent group.

- 30 As discussed above, using a method according to the above-described aspects of the invention it has now been discovered, most surprisingly, that the response to hypoxia differs between different specialised cell types or between different physiological states of the same cell type. For example, it has been found that in macrophage cells, different polypeptides are induced/repressed during different physiological states.

Furthermore, it has been found that a subset of this group of polypeptides are regulated only in activated macrophage cells. Macrophages possess various biological activities, including cytotoxic effects towards tumour cells and phagocytosis of bacteria or cellular debris. These form an important and potent arm of innate immunity, and as such must be finely regulated. In the absence of interactions with pathogens or other immune cells, the aforementioned activities of the macrophage are greatly reduced (i.e. resting macrophages). When given appropriate stimuli, such as contact with the lipopolysaccharide surface of bacteria, and/or exposure to T-cell derived interferon gamma, the functional activities of the macrophage are greatly potentiated (i.e. activated macrophage).

The expression of a further subset of these polypeptides has been found herein to be induced in activated macrophages under conditions of hypoxia, whilst a still further subset has been found herein to be repressed in activated macrophages under conditions of hypoxia.

In resting macrophage cells, it has been found that different polypeptides are induced/repressed during the biological response to hypoxia. For example, it has been found that a subset of this group of polypeptides are regulated only in resting macrophage cells. The expression of a further subset of these polypeptides has been found herein to be induced in resting macrophages under conditions of hypoxia, whilst a still further subset has been found herein to be repressed in resting macrophages under conditions of hypoxia.

According to a further aspect of the invention, there is provided a purified and isolated nucleic acid molecule that encodes a polypeptide according to any one of the aspects of the invention discussed above. Such a nucleic acid molecule may consist of the nucleic acid sequence as recited in any one of SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216, or form a redundant equivalent or fragment thereof. This aspect of the invention also includes a purified nucleic acid molecule which hybridizes under high stringency conditions with a nucleic acid molecule as described above.

According to a further aspect of the invention, there is provided an expression vector that contains a purified and isolated nucleic acid molecule according to the aspects of the invention described above. The invention also incorporates a delivery vehicle, such as a liposome, comprising a nucleic acid according to the above-described aspects of the invention.

In a further aspect, the invention provides a host cell transformed with a vector of the above-described aspect of the invention.

In a still further aspect, the invention provides a ligand that binds specifically to a polypeptide according to the above-described aspects of the invention. The ligand may be an antagonist ligand that inhibits the biological activity of the polypeptide, or may be an agonist ligand that activates the hypoxia-induced activity of the polypeptide to augment or potentiate a hypoxia-induced activity.

- 5 In a still further aspect of the invention, there is provided a ligand which binds specifically to, and which preferably inhibits the hypoxia-induced activity of, a polypeptide according to any one of the above-described aspects of the invention. Such a ligand may, for example, be an antibody that is immunospecific for the polypeptide in question.

According to a further aspect, the invention provides a polypeptide, a nucleic acid molecule, vector or
10 ligand as described above, for use in therapy or diagnosis of a disease or abnormal physiological condition. Preferably, the disease or abnormal physiological condition that is affected by hypoxia; examples of such diseases include cancer, ischaemic conditions (such as stroke, coronary arterial disease, peripheral arterial disease), reperfusion injury, retinopathy, neonatal stress, preeclampsia, atherosclerosis, inflammatory conditions (including rheumatoid arthritis), hair loss and wound healing. The undesired
15 cellular process involved in said diseases might include, but is not restricted to; tumorigenesis, angiogenesis, apoptosis, inflammation or erythropoiesis. The undesired biochemical processes involved in said cellular processes might include, but is not restricted to, glycolysis, gluconeogenesis, glucose transportation, catecholamine synthesis, iron transport or nitric oxide synthesis.

According to the invention, a number of known proteins have also been implicated in the biological
20 response to hypoxia. The functions of these proteins are known, meaning that these functions have been annotated in the public databases. The sequences of these proteins are presented in SEQ ID Nos.: 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 239, 241, 243, 245, 247, 249, 251, 253, 255, 257, 259, 261, 263, 265, 267, 269, 271, 273, 275, 277, 279, 281, 283, 285, 287, 289, 291, 293, 295, 297, 299, 301, 303, 305, 307, 309, 311, 313, 315, 317, 319, 321, 323, 325, 327, 329, 331, 333, 335, 337, 339, 341, 343,
25 345, 347, 349, 351, 353, 355, 357, 359, 361, 363, 365, 367, 369, 371, 373, 375, 377, 379, 381, 383, 385, 387, 389, 391, 393, 395, 397, 399, 401, 403, 405, 407, 409, 411, 413, 415, 417, 419, 421, 423, 425, 427, 429, 431, 433, 435, 437, 439, 441, 443, 445, 447, 449, 451, 453, 455, 457, 459, 461, 463, 465, 467, 469, 471, 473, 475, 477, 479, 481, 483, 485 and 487.

According to a further aspect of the invention, there is provided a substantially purified polypeptide,
30 which polypeptide:

- i) comprises the amino acid sequence as recited in any one of SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 63, 67, 69, 73, 75, 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129,

131, 133, 137, 139, 141, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 or
 209 or any one of SEQ ID Nos.: 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237,
 239, 241, 243, 245, 247, 249, 251, 253, 255, 257, 259, 261, 263, 265, 267, 269, 271,
 273, 275, 277, 279, 281, 283, 285, 287, 289, 291, 293, 295, 297, 299, 301, 303, 305,
 5 307, 309, 311, 313, 315, 317, 319, 321, 323, 325, 327, 329, 331, 333, 335, 337, 339,
 341, 343, 345, 347, 349, 351, 353, 355, 357, 359, 361, 363, 365, 367, 369, 371, 373,
 375, 377, 379, 381, 383, 385, 387, 389, 391, 393, 395, 397, 399, 401, 403, 405, 407,
 409, 411, 413, 415, 417, 419, 421, 423, 425, 427, 429, 431, 433, 435, 437, 439, 441,
 443, 445, 447, 449, 451, 453, 455, 457, 459, 461, 463, 465, 467, 469, 471, 473, 475,
 10 477, 479, 481, 483, 485, 487, 489 and 491;

ii) has an amino acid sequence encoded by a nucleic acid sequence recited in any one of
 SEQ ID Nos: 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96,
 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132,
 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166,
 15 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200,
 202, 204, 206, 208, 210, 212, 214 and 216, or encoded by a gene identified from an EST
 recited in any one of these SEQ ID Nos;

iii) is a fragment of a polypeptide according to i) or ii), provided that said fragment retains a
 biological activity possessed by the full length polypeptide of i) or ii), or has an
 20 antigenic determinant in common with the polypeptide of i) or ii); or

iv) is a functional equivalent of a polypeptide of i), ii) or (iii);

for use in the diagnosis or therapy of tumourigenesis, angiogenesis, apoptosis, the biological response to
 hypoxia conditions, or a hypoxic-associated pathology.

The invention also provides a purified and isolated nucleic acid molecule that encodes a polypeptide
 25 according to this aspect of the invention, for use in the diagnosis or therapy of tumourigenesis,
 angiogenesis, apoptosis, the biological response to hypoxia conditions, or a hypoxic-associated pathology.

The sequences of these molecules are provided in SEQ ID Nos.: 218, 220, 222, 224, 226, 228, 230, 232,
 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, 268, 270, 272, 274,
 276, 278, 280, 282, 284, 286, 288, 290, 292, 294, 296, 298, 300, 302, 304, 306, 308, 310, 312, 314, 316,
 30 318, 320, 322, 324, 326, 328, 330, 332, 334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358,
 360, 362, 364, 366, 368, 370, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 392, 394, 396, 398, 400,
 402, 404, 406, 408, 410, 412, 414, 416, 418, 420, 422, 424, 426, 428, 430, 432, 434, 436, 438, 440, 442,
 444, 446, 448, 450, 452, 454, 456, 458, 460, 462, 464, 466, 468, 470, 472, 474, 476, 478, 480, 482, 484,

486 and 488. As described above for the EST nucleic acid sequences annotated herein, this aspect of the invention includes redundant equivalents and fragments of the sequences explicitly recited in SEQ ID Nos.: 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, 268, 270, 272, 274, 276, 278, 280, 282, 284, 286, 288, 290, 292, 294, 296, 298, 5 300, 302, 304, 306, 308, 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330, 332, 334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 370, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 392, 394, 396, 398, 400, 402, 404, 406, 408, 410, 412, 414, 416, 418, 420, 422, 424, 426, 428, 430, 432, 434, 436, 438, 440, 442, 444, 446, 448, 450, 452, 454, 456, 458, 460, 462, 464, 466, 468, 470, 472, 474, 476, 478, 480, 482, 484, 486 and 488, and purified nucleic acid molecules which 10 hybridize under high stringency conditions with such nucleic acid molecules, and vectors containing such nucleic acid molecules for use in the diagnosis or therapy of tumourigenesis, angiogenesis, apoptosis, the biological response to hypoxia conditions, or a hypoxic-associated pathology.

This aspect of the invention also includes ligands which bind specifically to, and which preferably inhibit the hypoxia-induced activity of, a polypeptide listed in SEQ ID Nos.: 217, 219, 221, 223, 225, 227, 229, 15 231, 233, 235, 237, 239, 241, 243, 245, 247, 249, 251, 253, 255, 257, 259, 261, 263, 265, 267, 269, 271, 273, 275, 277, 279, 281, 283, 285, 287, 289, 291, 293, 295, 297, 299, 301, 303, 305, 307, 309, 311, 313, 315, 317, 319, 321, 323, 325, 327, 329, 331, 333, 335, 337, 339, 341, 343, 345, 347, 349, 351, 353, 355, 357, 359, 361, 363, 365, 367, 369, 371, 373, 375, 377, 379, 381, 383, 385, 387, 389, 391, 393, 395, 397, 399, 401, 403, 405, 407, 409, 411, 413, 415, 417, 419, 421, 423, 425, 427, 429, 431, 433, 435, 437, 439, 20 441, 443, 445, 447, 449, 451, 453, 455, 457, 459, 461, 463, 465, 467, 469, 471, 473, 475, 477, 479, 481, 483, 485 and 487, for use in the diagnosis or therapy of tumourigenesis, angiogenesis, apoptosis, the biological response to hypoxia conditions, or a hypoxic-associated pathology.

The invention also provides a pharmaceutical composition suitable for modulating hypoxia and/or ischaemia, comprising a therapeutically-effective amount of a a polypeptide, a nucleic acid molecule, 25 vector or ligand as described above, in conjunction with a pharmaceutically-acceptable carrier.

The invention also provides a vaccine composition comprising a polypeptide, or a nucleic acid molecule as described above.

The invention also provides a method of treating a disease in a patient in need of such treatment by administering to a patient a therapeutically effective amount of a polypeptide, a nucleic acid molecule, 30 vector, ligand or pharmaceutical composition as described above. For diseases in which the expression of the natural gene or the activity of the polypeptide is lower in a diseased patient when compared to the level of expression or activity in a healthy patient, the polypeptide, nucleic acid molecule, ligand, compound or composition administered to the patient should be an agonist. For diseases in which the

expression of the natural gene or activity of the polypeptide is higher in a diseased patient when compared to the level of expression or activity in a healthy patient, the polypeptide, nucleic acid molecule, vector, ligand, compound or composition administered to the patient is an antagonist. By the term "agonist" is meant herein, any polypeptide, peptide, synthetic molecule or organic molecule that functions as an activator, by increasing the effective biological activity of a polypeptide, for example, by increasing gene expression or enzymatic activity. By the term "antagonist" is meant herein, any polypeptide, peptide, synthetic molecule or organic molecule that functions as an inhibitor, by decreasing the effective biological activity of the gene product, for example, by inhibiting gene expression of an enzyme or a pharmacological receptor.

- 10 The invention also provides for the use of a polypeptide, nucleic acid molecule, vector, ligand or pharmaceutical composition according to any one of the above-described aspects of the invention in modifying the response of a cell to conditions of hypoxia.

The invention also provides a polypeptide, nucleic acid molecule, vector, ligand or pharmaceutical composition according to any one of the above-described aspects of the invention, for use in the manufacture of a medicament for the treatment of a hypoxia-regulated condition.

The invention also provides a method of monitoring the therapeutic treatment of disease or physiological condition in a patient, comprising monitoring over a period of time the level of expression or activity of polypeptide, nucleic acid molecule, vector or ligand in tissue from said patient, wherein altering said level of expression or activity over the period of time towards a control level is indicative of regression of said disease or physiological condition.

The invention also provides a method of providing a hypoxia regulating gene, an apoptotic or an angiogenesis regulating gene by administering directly to a patient in need of such therapy an expressible vector comprising expression control sequences operably linked to one or more of the nucleic acid molecules as described above.

- 25 The invention also provides a method of diagnosing a hypoxia-regulated condition in a patient, comprising assessing the level of expression of a natural gene encoding a polypeptide according to any one of the aspects of the invention described above in tissue from said patient and comparing said level of expression or activity to a control level, wherein a level that is different to said control level is indicative of the hypoxia-related condition.

- 30 Such a method of diagnosis may be carried out *in vitro*. One example of a suitable method comprises the steps of: (a) contacting a ligand as described above with a biological sample under conditions suitable for the formation of a ligand-polypeptide complex; and (b) detecting said complex.

A further example of a suitable method may comprises the steps of: a) contacting a sample of tissue from the patient with a nucleic acid probe under stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule whose sequence is recited in any one of SEQ ID Nos.: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, 268, 270, 272, 274, 276, 278, 280, 282, 284, 286, 288, 290, 292, 294, 296, 298, 300, 302, 304, 306, 308, 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330, 332, 334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 370, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 392, 394, 396, 398, 400, 402, 404, 406, 408, 410, 412, 414, 416, 418, 420, 422, 424, 426, 428, 430, 432, 434, 436, 438, 440, 442, 444, 446, 448, 450, 452, 454, 456, 458, 460, 462, 464, 466, 468, 470, 472, 474, 476, 478, 480, 482, 484, 486 and 488, and the probe; b) contacting a control sample with said probe under the same conditions used in step a); and c) detecting the presence of hybrid complexes in said samples; wherein detection of levels of the hybrid complex in the patient sample that differ from levels of the hybrid complex in the control sample is indicative of the hypoxia-related condition.

A still further example of a suitable method may comprise the steps of: a) contacting a sample of nucleic acid from tissue of the patient with a nucleic acid primer under stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule whose sequence is recited in any one of SEQ ID Nos.: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, 268, 270, 272, 274, 276, 278, 280, 282, 284, 286, 288, 290, 292, 294, 296, 298, 300, 302, 304, 306, 308, 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330, 332, 334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 370, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 392, 394, 396, 398, 400, 402, 404, 406, 408, 410, 412, 414, 416, 418, 420, 422, 424, 426, 428, 430, 432, 434, 436, 438, 440, 442, 444, 446, 448, 450, 452, 454, 456, 458, 460, 462, 464, 466, 468, 470, 472, 474, 476, 478, 480, 482, 484, 486 and 488, and the primer; b) contacting a control sample with said primer under the same conditions used in step a); c) amplifying the sampled nucleic acid; and d) detecting the level of

amplified nucleic acid from both patient and control samples; wherein detection of levels of the amplified nucleic acid in the patient sample that differ significantly from levels of the amplified nucleic acid in the control sample is indicative of the hypoxia-related condition.

A still further example of a suitable method may comprised the steps of: a) obtaining a tissue sample from
5 a patient being tested for the hypoxia-related condition; b) isolating a nucleic acid molecule according to any one of the above-described aspects of the invention from said tissue sample; and c) diagnosing the patient for the hypoxia-related condition by detecting the presence of a mutation which is associated with the hypoxia-related condition in the nucleic acid molecule as an indication of the hypoxia-related condition. This method may comprise the additional step of amplifying the nucleic acid molecule to form
10 an amplified product and detecting the presence or absence of a mutation in the amplified product.

Particular hypoxia-related conditions that may be diagnosed in this fashion include cancer, ischaemia, reperfusion, retinopathy, neonatal stress, preeclampsia, atherosclerosis, rheumatoid arthritis, undesired hair loss, cardiac arrest or stroke, for example, caused by a disorder of the cerebral, coronary or peripheral circulation.

15 In a further aspect, the invention provides a method for the identification of a compound that is effective in the treatment and/or diagnosis of a hypoxia-regulated condition, comprising contacting a polypeptide, nucleic acid molecule, or ligand according to any one of the above-described aspects of the invention with one or more compounds suspected of possessing binding affinity for said polypeptide, nucleic acid molecule or ligand, and selecting a compound that binds specifically to said nucleic acid molecule,
20 polypeptide or ligand.

According to a still further aspect of the invention, there is provided a kit useful for diagnosing a hypoxia-regulated condition, comprising a first container containing a nucleic acid probe that hybridises under stringent conditions with a nucleic acid molecule according to any one of the aspects of the invention described above; a second container containing primers useful for amplifying said nucleic acid molecule;
25 and instructions for using the probe and primers for facilitating the diagnosis of the hypoxia-regulated condition. The kit may additionally comprise a third container holding an agent for digesting unhybridised RNA.

To facilitate in the diagnosis of the hypoxia-regulated condition using one of the methods outlined above, in a further aspect, the invention provides an array of at least two nucleic acid molecules, wherein each of
30 said nucleic acid molecules either corresponds to the sequence of, is complementary to the sequence of, or hybridises specifically to a nucleic acid molecule according to any one of the aspects of the invention described above. Such an array may contain nucleic acid molecules that either correspond to the sequence of, are complementary to the sequence of, or hybridise specifically to at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11,

12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 92a, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 5 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 10 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295 or more of the nucleic acid molecules implicated in a hypoxia-regulated condition as recited above. The nucleic acid molecules on the array 15 may consist of oligonucleotides of between twelve and fifty nucleotides, more preferably, between forty and fifty nucleotides. Alternatively, the nucleic acid molecules on the array may consist of PCR-amplified cDNA inserts where the nucleic acid molecule is between 300-2000 nucleotides.

In a related aspect, again useful for diagnosis, the invention provides an array of antibodies, comprising at least two different antibody species, wherein each antibody species is immunospecific with a polypeptide 20 implicated in a hypoxia-regulated condition as described above. The invention also provides an array of polypeptides, comprising at least two polypeptide species as recited above, wherein each polypeptide species is implicated in a hypoxia-regulated condition, or is a functional equivalent variant or fragment thereof.

Kits useful in the diagnostic methods of the invention may comprise such nucleic acid, antibody and/or 25 polypeptide arrays.

According to the invention, a kit may also comprise one or more antibodies that bind to a polypeptide as recited above, and a reagent useful for the detection of a binding reaction between said antibody and said polypeptide.

According to a still further aspect of the invention, there is provided a genetically-modified non-human 30 animal that has been transformed to express higher, lower or absent levels of a polypeptide according to any one of the aspects of the invention described above. Preferably, said genetically-modified animal is a transgenic or knockout animal.

The invention also provides a method for screening for a compound effective to treat a hypoxia-regulated condition, by contacting a non-human genetically-modified animal as described above with a candidate compound and determining the effect of the compound on the physiological state of the animal.

As discussed in some detail above, ischaemic disease pathologies involve a decrease in the blood supply to a bodily organ, tissue or body part generally caused by constriction or obstruction of the blood vessels. One particular example of an ischaemic disease pathology is myocardial ischaemia, which encompasses several chronic and acute cardiac pathologies that involve the deprivation of the myocardium of its blood supply, usually through coronary artery occlusion. A key component of ischaemia is hypoxia. Following transient ischaemia, the affected tissue may be subjected to reperfusion and re-oxygenation, and this is of significance in its own right.

Ischaemia/reperfusion is well known to induce cell death in myocardial tissue by apoptosis, leading to impaired function of the myocardium and infarction. Many of the specific molecules required to execute the process of apoptosis are known, but not all of these molecules have been characterised in detail. Cell death may also proceed by a distinct process called necrosis, which unlike apoptosis, is not initiated and controlled by specific and dedicated cellular and biochemical mechanisms (see Nicotera *et al.*, Biochem Soc Symp. 1999; 66:69-73). There is substantial evidence that apoptotic cell death occurs either during or after myocardial ischaemia (Kajstura *et al.*, Lab Invest. 1996; 74(1):86-107; Cheng *et al.*, Exp Cell Res. 1996; 226(2):316-27; Fliss and Gattinger, Circ Res. 1996; 79(5):949-56; Veinot *et al.*, Hum Pathol. 1997; 28(4):485-92; Bialik *et al.*, J Clin Invest. 1997; 100(6):1363-72; Gottlieb *et al.*, J Clin Invest. 1994; 94(4):1621-8; Gottlieb and Engler, Ann N Y Acad Sci. 1999; 874:412-26). In the laboratory, apoptosis is also induced by subjecting cardiac myocytes to hypoxia (Tanaka *et al.*, Circ Res. 1994 Sep;75(3):426-33; Long *et al.*, J Clin Invest. 1997 99(11): 2635-43).

Clearly, there is a significant clinical application were a successful method to inhibit apoptosis in ischaemic myocardial tissue to be devised. A specific and effective treatment requires identifying biochemical target(s), which are responsible for mediating apoptosis, specifically in ischaemic myocardial cells. One target which plays a common role in mediating apoptosis in many cell types, namely p53, is not involved in apoptosis resulting from myocardial ischaemia (Bialik *et al.*, J Clin Invest. 1997; 100(6):1363-72). Others have shown that inhibiting key mediators of apoptosis, caspases, provides protection against lethal reperfusion injury, following myocardial ischaemia in rat models (Mocanu *et al.*, Br J Pharmacol. 2000; 130(2):197-200; Yaoita *et al.*, Circulation. 1998 97(3): 276-81; Holly *et al.*, J Mol Cell Cardiol. 1999 31(9): 1709-15). However, this approach lacks specificity, since the caspases play a key role in mediating apoptosis in the majority of mammalian cell types, where it is usually beneficial. An

approach that involves modulating the activity of molecules shown specifically to mediate apoptosis in ischaemic cardiac cells, would present a distinct advantage in both specificity and efficacy.

It has now been discovered that a polypeptide encoded by a gene identified from the EST recited in SEQ ID No 86, having the Protein accession number BAB15101 (encoded by Homo sapiens cDNA: FLJ21620
 5 fis, clone COL07838 Nucleotide accession AK025273) is regulated by hypoxia. Other public domain sequences corresponding to this gene include Homo sapiens cDNA: FLJ23265 fis, clone COL06456 Nucleotide accession AK026918. Accordingly, when referring in the present specification to the EST recited in SEQ ID No 86, it is intended that these gene and protein sequences are also embraced. This gene was identified using Research Genetics Human GeneFilters arrays, which contain an EST
 10 corresponding to the gene (accession number R00332). In the art, the gene is now termed EGL nine (C.elegans) homolog 3.

There are no reports that describe the function of this human gene. However, a high degree of amino acid homology is observed between the protein encoded by this gene, and a rat protein called "Growth factor responsive smooth muscle protein" or "SM20" (Nucleotide accession U06713; Protein accession
 15 A53770). An alignment of single letter amino acid sequences is shown below. Over the highlighted region there is 97% amino acid similarity and 96% amino acid identity.

	A53770	(1)	MTLRSRRGFLSFLPGLRPPRRWLRI SKRGPPTSHWASPALGGRTLHYSCR	
	BAB15101	(1)	-----	
20			51	100
	A53770	(51)	SQSGTPFSSEFQATFPFAAKVARGPWLPQVVEPPARLSASPLCVRSGQA	
	BAB15101	(1)	-----	
			101	150
	A53770	(101)	LGACTLGVPRLGSVSEMP LGHIMRLDLEKIALEYIVPCLHEVGFCYLDNF	
25	BAB15101	(1)	-----MPLGHIMRLDLEKIALEYIVPCLHEVGFCYLDNF	
			151	200
	A53770	(151)	LGEVVGDCVLERVKQLHYNGALRDGQLAGPRAGVSKRHLRGDQITWIGGN	
	BAB15101	(35)	LGEVVGDCVLERVKQLHCTGALRDGQLAGPRAGVSKRHLRGDQITWIGGN	
			201	250
30	A53770	(201)	EEGCEAINFLLSLIDRLVLYCGSRLGKYYVKERSKAMVACYPGNGTGYVR	
	BAB15101	(85)	EEGCEAISFLLSLIDRLVLYCGSRLGKYYVKERSKAMVACYPGNGTGYVR	
			251	300
	A53770	(251)	HVDNPNGDGRCITCIYYLNKNWDAKLHGGVLRIFPEGKSFVADVEPIFDR	
35	BAB15101	(135)	HVDNPNGDGRCITCIYYLNKNWDAKLHGGILRIFPEGKSFVADVEPIFDR	
			301	350
	A53770	(301)	LLFSWSDRRNPHEVQPSYATRYAMTVWYFDAEERAEAKKKFRNLTRKTES	
	BAB15101	(185)	LLFFWSDRRNPHEVQPSYATRYAMTVWYFDAEERAEAKKKFRNLTRKTES	
			351	
	A53770	(351)	ALAKD	
40	BAB15101	(235)	ALTED	

The high degree of amino acid similarity suggests that the human protein BAB15101 has an equivalent biochemical function to the rat protein A53770 ("Growth factor responsive smooth muscle protein" or "SM20"). Recent publications have shown that SM20 functions to promote apoptosis in neurons (Lipscomb *et al.*, *J Neurochem* 1999; 73(1):429-32; Lipscomb *et al.*, *J Biol Chem* 2000 Nov 1; [epub
5 ahead of print]). Significantly, SM20 has been shown to be expressed at high levels in the heart (Wax *et al.*, *J Biol Chem* 1994; 269(17): 13041-7).

It has also been discovered that a polypeptide encoded by a gene identified from the EST recited in SEQ ID No 90, having the Protein accession number CAB81622, is regulated by hypoxia. The encoding human gene has been annotated in the UniGene database as "Similar to rat smooth muscle protein SM-
10 20"; the nucleotide sequence is contained within the nucleotide accession AL117352. More recently, a longer fragment of this gene has been cloned, named clorf12, or EGLN1 (Nucleotide accession AAG34568; Protein accession AAG34568). Accordingly, when referring in the present specification to the EST recited in SEQ ID No 90, it is intended that these gene and protein sequences are also embraced.

This distinct human gene, encoding a protein related to SM20 and EGLN3 (BAB15101), is also induced
15 in response to hypoxia. This gene was identified using Research Genetics Human GeneFilters arrays, which contain an EST corresponding to the gene (accession number H56028).

Independently to this, a fragment of this gene has been cloned from a cDNA library derived from hypoxic human cardiomyoblasts, and it has been shown that the gene is increased in expression in response to hypoxia in this cell type (see Table I herein; penultimate row). The nucleotide sequence of this cDNA
20 fragment is referred to herein as SEQ ID No 90a.

In the light of this novel discovery reported herein that these human equivalents of SM20 are induced by hypoxia, it is herein proposed that in cardiac ischaemia, the resulting apoptosis is due at least in part, to increased expression of these genes.

The therapeutic modulation of the activity of EGLN3 (BAB15101), clorf12 (AAG34568), CAB81622,
25 SM20 and other equivalent proteins and encoding genes therefore provides a novel means for the treatment of myocardial ischaemia, through the alteration of the propensity of myocardial cells to undergo apoptosis. For example, a suitable treatment may involve altering the susceptibility of ischaemic myocardial tissue to subsequent reperfusion and re-oxygenation, or may involve modulating the susceptibility of chronic ischaemic myocardial tissue (including forms of angina) to later more severe
30 ischaemia, which would result in myocardial infarction. It is submitted that, by way of analogy, cerebral ischaemia may be treated using the same principle.

These data provide the first connection between these related genes and the physiological response to hypoxia. Recently published research papers have identified that the protein products of these genes can

act as proline hydroxylases (see Bruick RK et al Science. 2001 294:1337-40 and Epstein AC et al Cell. 107:43-54). This is consistent with our observations that certain proline hydroxylases are induced in response to hypoxia and the genes EGLN1 and EGLN3 are part of the hypoxia response. For example, two genes encoding proline hydroxylases have been identified herein as being increased in expression in response to hypoxia (proline 4-hydroxylase, alpha polypeptide I; SeqID: 231/232, proline 4-hydroxylase, alpha polypeptide II; SeqID: 349/ 350). This identified a functional significance of proline hydroxylation as a response to hypoxia. A preferred embodiment of the invention thus includes methods for modulating the biological response to hypoxia by modulating the proline hydroxylase activity of the EGLN3 (BAB15101), c1orf12 (AAG34568), CAB81622 and SM20 proteins.

Furthermore, a number of bacteria, such as *moraxella*, are thought to be involved in the initiation of inflammatory diseases. Many bacteria contain, within their genome, genes encoding proteins that share homology to the EGLN family of prolyl hydroxylases. We therefore propose that these bacterial genes may initiate a hypoxic like response at the site of infection thereby causing localised inflammation. The resulting inflammatory infiltrate could then cause the tissue to become hypoxic thereby continuing the cycle of hypoxia response.

As discussed in detail above, fragments and functional equivalents of the EGLN3 (BAB15101), c1orf12 (AAG34568), CAB81622, SM20 and other equivalent proteins are included within the present invention, in addition to ligands that bind specifically to these proteins. Furthermore, the invention also embraces purified and isolated nucleic acid molecules encoding these proteins, fragments and functional equivalents, vectors containing such nucleic acid molecules and host cells transformed with these vectors.

The therapeutic and diagnostic applications discussed above are also equally relevant to this aspect of the invention. For example, small molecule inhibitors of the EGLN3 (BAB15101), c1orf12 (AAG34568), CAB81622, SM20 and equivalent proteins and encoding genes are envisaged for utility as pharmaceutical agents, particularly in modulating the proline hydroxylase activity of the EGLN3 and c1orf12 proteins.

Truncated or chimeric inhibitory derivatives of the encoding genes, or distinct genes that encode regulators of the BAB15101, AAG34568, CAB81622 and SM20 encoding genes, are also envisaged for utility for gene therapy.

An alignment of the amino acid sequences of rat SM20 (Accession A53770), its human equivalent (Accession BAB15101; SEQ ID No: 85) and this distinct human homologue (Accession CAB81622 or AAG34568; SEQ ID No: 89) is shown below:

		1		50
	BAB15101	(1)	-----	
	A53770	(1)	-----	
35	AAG34568	(1)	MANDSGGPGGPPSPSERDRQYCELCGKMENLLRCSRCSFFYCCKEHQRQD	

Consensus	(1)	51	100
BAB15101	(1)	-----	
A53770	(1)	-----MTLRSRRGFLSFLPGLRPPRRWLRISKRGPPPTSHWASP-----AL	
5 AAG34568	(51)	WKKHKLVLCQSGSEGALGHGVGPHQHS GPAPPA AVPPPPRAGAREPRKAAARR	
Consensus	(51)	L G L G A P P A P	
		101	150
BAB15101	(1)	-----	
A53770	(41)	GGRTLHYSCRSQSGTPFSSEFQATFPAPAFAAKVARGPWLPQVVEPPAR---	
10 AAG34568	(101)	DNASGDAAKGKVKAKPPADPAAAAAPCAAAGGQGSAAVAAEAEPGKEEPP	
Consensus	(101)	S A A P A A P AA A G L EP	
		151	200
BAB15101	(1)	-----MPLGHIMRLDLEKIALEYIVP	
A53770	(88)	LSASPLCVRSGQALGACTLGVPRLGSVSEMP LGHIMRLDLEKIALEYIVP	
15 AAG34568	(151)	ARSSLFOEKANLYPPSNTPGDALSPGGGLRPNQGTKPLPALKLALEYIVP	
Consensus	(151)	AS KA A T G MPLGHIMRLDLEKIALEYIVP	
		201	250
BAB15101	(22)	CLHEVGFCYLDNFLGEVVGDCVLERVKQLHCTGALRDGQLAGPRAGVSKR	
A53770	(138)	CLHEVGFCYLDNFLGEVVGDCVLERVKQLHYNGALRDGQLAGPRAGVSKR	
20 AAG34568	(201)	CMNKHGICVDDFLGKETGOQIGDEVRAHDTGKFTDGQLVSQKS-DSSK	
Consensus	(201)	CLHEVGFCYLDNFLGEVVGDCVLERVKQLH TGALRDGQLAGPRAGVSKR	
		251	300
BAB15101	(72)	HLRGDQITWIGGNEEGCEATSFLLSLIDRLVLYCGSRLGKYYVKERSKAM	
A53770	(188)	HLRGDQITWIGGNEEGCEATNFFLLSLIDRLVLYCGSRLGKYYVKERSKAM	
25 AAG34568	(250)	DIRGDKITWIEGKEPCETIGLLMSSMDDLIRHCNGKLGSKYKINGRTKAM	
Consensus	(251)	HLRGDQITWIGGNEEGCEAI FLLSLIDRLVLYCGSRLGKYYVKERSKAM	
		301	350
BAB15101	(122)	VACYPGNGTGYVRHVDNPNGDGRCITCIYYLNKNWDAKLHGGILRIFPEG	
A53770	(238)	VACYPGNGTGYVRHVDNPNGDGRCITCIYYLNKNWDAKLHGGVLRIFPEG	
30 AAG34568	(300)	VACYPGNGTGYVRHVDNPNGDGRCVTCTIYYLNKNWDAKVS GGILRIFPEG	
Consensus	(301)	VACYPGNGTGYVRHVDNPNGDGRCITCIYYLNKNWDAKLHGGILRIFPEG	
		351	400
BAB15101	(172)	KSFADVEPIFDRLFFWSDRRNPEHVQPSYATRYAMTVWYFDAEERAEAE	
A53770	(288)	KSFVADVEPIFDRLFFWSDRRNPEHVQPSYATRYAMTVWYFDAEERAEAE	
35 AAG34568	(350)	KAQFADIEPKFDRLFFWSDRRNPEHVQPAVATRYAITVWYFDADERARA	
Consensus	(351)	KSFADVEPIFDRLFFWSDRRNPEHVQPSYATRYAMTVWYFDAEERAEAE	
		401	427
BAB15101	(222)	KKKFRNLTRKTESALTED-----	
A53770	(338)	KKKFRNLTRKTESALAKD-----	
40 AAG34568	(400)	KVKYLTGEKGVRELNKPSDSVSGKDVF	
Consensus	(401)	KKKFRNLTRKTESAL KD	

From this sequence alignment, a highly conserved region of amino acid sequence may be noted, the consensus of which is as follows:

45

KAMVACYPGNGTGYVRHVDNPNGDGRCITCIYYLNKNWDAKLHGGILRIFPEGKSFADVEPIFDRLFF
WSDRRNPHEVQPSYATRYAMTVWYFDAEERAEAKK

This consensus sequence, and variants thereof, may be used in the identification of other proteins that are implicated in the biological response to hypoxia. This aspect of the invention therefore provides a substantially purified polypeptide comprising the consensus sequence:

- 5 KAMVACYPGNGTGYVRHVDNPNNGDGRCITCIYYLNKNWDAKLHGGILRIFPEGKSFIADVEPIFDRLLFF
WSDRRNPHEVQPSYATRYAMTVWYFDAEERAEAKKK, or a variant thereof.

The invention also provides a substantially purified polypeptide comprising the consensus sequence:
KAMVACYPGNGTGYVRHVDNPNNGDGRCITCIYYLNKNWDAKLHGGILRIFPEGKSFIADVEPIFDRLLFF

- 10 WSDRRNPHEVQPSYATRYAMTVWYFDAEERAEAKKK, or a variant thereof, in the treatment or diagnosis
of a hypoxia-related disease or condition.

Neither this consensus domain nor any proteins that contain this domain have been previously associated with the cellular response to hypoxia/ischaemia. Searches of the public databases indicate that the human genome contains several genes that encode proteins that contain this consensus sequence. These proteins
15 may have similar functions or may function in the same biochemical pathway, potentially with an antagonistic effect.

- By "variant" is meant a variation of the consensus sequence given above, that exhibits a degree of homology with the consensus sequence above a certain threshold level of identity or similarity. Degrees of identity and similarity can be readily calculated according to methods known in the art (see, for
20 example, Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing. Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993). Typically, greater than 50% identity between two sequences is considered to be an indication of functional equivalence. Preferably, a variant consensus according to this aspect of the invention exhibits a degree of sequence identity with the consensus sequence given above, of greater than 50%. More
25 preferred polypeptides have degrees of identity of greater than 60%, 70%, 80%, 90%, 95%, 98% or 99%, respectively.

- As discussed in detail above, fragments and functional equivalents of these proteins are included within the present invention, in addition to ligands that bind specifically to these proteins. Furthermore, the invention also embraces purified and isolated nucleic acid molecules encoding these proteins, fragments
30 and functional equivalents, vectors containing such nucleic acid molecules and host cells transformed with these vectors. The therapeutic and diagnostic applications discussed above are also equally relevant to this aspect of the invention.

The polypeptide referred to above as that encoded by SEQ ID No 91 is a specific protein that is termed "Semaphorin 4b". The gene encoding this protein is regulated (activated) by conditions of hypoxia. The Semaphorin 4b protein is encoded by a gene identified from the EST recited in SEQ ID No 92. The unequivocal and accurate full length cDNA sequence is provided herein as SEQ ID No 92a. The accurate
5 presumptive amino acid sequence is provided herein as SEQ ID No 91. This protein, functionally-equivalent variants of this protein, the encoding nucleic acid molecules and ligands that regulate the activity and/or expression of this gene and protein are claimed above in the context of their role in hypoxia and hypoxia-related disorders.

Semaphorins are a large family of proteins, characterised by the 500 amino acid sema domain (Puschel et al., 1995, Neuron, 14(5): 941-8; Tamagnone and Comoglio, 2000, Trends Cell Biol., 10(9): 377-83).
10 Early work showed a role in the guidance of axons during brain development, and the regulation of cell migration. More recently, specific members of this large family have been associated with cancer (Brambilla et al., Am J Pathol., 2000, 156(3): 939-50), rheumatoid arthritis (Mangasser-Stephan et al., Biochem Biophys Res Commun., 1997, 234(1): 153-6), the immune system (Spriggs, Curr Opin
15 Immunol., 1999, 11(4): 387-91) including B-lymphocyte functions (Hall et al., Proc Natl Acad Sci U S A, 1996, 93(21): 11780-5) and angiogenesis (Miao et al., J Cell Biol., 1999, 146(1): 233-42). This is perhaps not surprising considering that cell migration / trafficking is a key part of inflammation, angiogenesis and tumour metastasis.

There are at least distinct 25 human semaphorin genes and the significance/ utility of many of these
20 remains untested. This includes the Semaphorin 4b protein, which is unpublished and until now has not been assigned a full and accurate amino acid sequence.

We have made experimental discoveries which link the expression of Semaphorin 4b to factors (hypoxia, gamma IFN and superoxide radicals) that are associated with a variety of human ischaemic and inflammatory diseases. In particular, a key response of cells to hypoxia is to stimulate angiogenesis, and a
25 key part of inflammation is the recruitment and trafficking of immune cells. In light of our discoveries, and what is known about other specific members of the semaphorin family, it is herein proposed that Semaphorin 4b is a regulator of these cellular functions, and thus provides a novel target for therapeutic intervention. This paves the way for the development of therapeutic agents that either potentiate or antagonise functions of Semaphorin 4b. Such agents are likely to be highly valuable in the treatment of
30 human disease.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA technology and immunology, which are within the skill of those working in the art.

Most general molecular biology, microbiology recombinant DNA technology and immunological techniques can be found in Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (1989) Cold Harbor-Laboratory Press, Cold Spring Harbor, N.Y. or Ausubel *et al.*, Current protocols in molecular biology (1990) John Wiley and Sons, N.Y.

- 5 Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs.

A. Polypeptides

The term "polypeptide" as used herein, refers to a chain (may be branched or unbranched) of two or more amino acids linked to each other by means of a peptide bond or modified peptide bond (isosteres). The
10 term polypeptide encompasses but is not limited to oligopeptides, peptides and proteins. The polypeptide of the invention may additionally be either in a mature protein form or in a pre-, pro- or prepro-protein form that requires subsequent cleavage for formation of the active mature protein. The pre-, pro-, prepro-part of the protein is often a leader or secretory sequence but may also be an additional sequence added to aid protein purification (for example, a His tag) or to conform a higher stability to the protein.

- 15 A polypeptide according to the invention may also include modified amino acids, that is, amino acids other than those 20 that are gene-encoded. This modification may be a result of natural processes such as post-translational processing or by chemical modification. Examples of modifications include acetylation, acylation, amidation, ADP-ribosylation, arginylation, attachment of a lipid derivative or phosphatidylinositol, γ -carboxylation, covalent attachment of a flavin or haeme moiety, a nucleotide or
20 nucleotide derivative, cyclisation, demethylation, disulphide bond formation, formation of covalent cross-links, formylation, glycosylation, GPI anchor formation, hydroxylation, iodination, lipid attachment, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemisation, selenoylation, sulphation, and ubiquitination. Modification of the polypeptide can occur anywhere within the molecule including the backbone, the amino acid side-chains or at the N- or C-
25 terminals.

A polypeptide according to the invention may either be isolated from natural sources (for example, purified from cell culture), or be a recombinantly produced polypeptide, or a synthetically produced polypeptide or a combination of all the above.

Antibodies

- 30 A polypeptide according to the invention, its functional equivalents and/or any immunogenic fragments derived from the polypeptide may be used to generate ligands including immunospecific monoclonal or polyclonal antibodies, or antibody fragments. These antibodies can then be used to isolate or identify

clones expressing the polypeptide of the invention or to purify the polypeptide by affinity chromatography. Further uses of these immunospecific antibodies may include, but are not limited to, diagnostic, therapeutic or general assay applications. Examples of assay techniques that employ antibodies are immunoassays, radioimmunoassays (RIA) or enzyme linked immunosorbent assay
5 (ELISA). In these cases, the antibodies may be labelled with an analytically-detectable reagent including radioisotopes, a fluorescent molecule or any reporter molecule.

The term "immunospecific" as used herein refers to antibodies that have a substantially higher affinity for a polypeptide of this invention compared with other polypeptides. The term "antibody" as used herein refers to a molecule that is produced by animals in response to an antigen and has the particular property
10 of interacting specifically with the antigenic determinant that induced its formation. Fragments of the aforementioned molecule such as Fab, F(ab')₂ and scFv, which are capable of binding the antigen determinant, are also included in the term "antibody". Antibodies may also be modified to make chimeric antibodies, where non-human variable regions are joined or fused to human constant regions (for example, Liu *et al.*, PNAS, USA, 84, 3439 (1987)). Particularly, antibodies may be modified to make
15 them less immunogenic to an individual in a process such as humanisation (see, for example, Jones *et al.*, Nature, 321, 522 (1986); Verhoeven *et al.*, Science, 239, 1534 (1988); Kabat *et al.*, J. Immunol., 147, 1709 (1991); Queen *et al.*, PNAS, USA, 86, 10029 (1989); Gorman *et al.*, PNAS, USA, 88, 34181 (1991) and Hodgson *et al.*, Bio/Technology, 9, 421 (1991)). The term "humanised antibody", as used herein, refers to antibody molecules in which the amino acids of the CDR (complementarity-determining region)
20 and selected other regions in the variable domains of the heavy and/or light chains of a non-human donor antibody have been substituted with the equivalent amino acids of a human antibody. The humanised antibody therefore closely resembles a human antibody, but has the binding ability of the donor antibody. Antibodies may also have a "bispecific" nature, that is, the antibody has two different antigen binding domains, each domain being directed against a different epitope.

25 Specific polyclonal antibodies may be made by immuno-challenging an animal with a polypeptide of this invention. Common animals used for the production of antibodies include the mouse, rat, chicken, rabbit, goat and horse. The polypeptide used to immuno-challenge the animal may be derived by recombinant DNA technology or may be chemically-synthesised. In addition, the polypeptide may be conjugated to a carrier protein. Commonly used carriers to which the polypeptides may be conjugated include, but are not
30 limited to BSA (bovine serum albumin), thyroglobulin and keyhole limpet haemocyanin. Serum from the immuno-challenged animal is collected and treated according to known procedures, for example, by immunoaffinity chromatography.

Specific monoclonal antibodies can generally be made by methods known to one skilled in the art (see for

example, Kohler, G. and Milstein, C., *Nature* 256, 495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, 77-96 in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985) and Roitt, I. *et al.*, *Immunology*, 25.10, Mosby-Year Book Europe Limited (1993)). Panels of monoclonal antibodies produced against the polypeptides of the invention can be screened for various properties, i.e.,
5 for isotype, epitope, affinity, etc. against which they are directed. Alternatively, genes encoding the monoclonal antibodies of interest may be isolated from hybridomas, for instance using PCR techniques known in the art, and cloned and expressed in appropriate vectors.

Phage display technology may be utilised to select the genes encoding the antibodies that have exhibited an immunespecific response to the polypeptides of the invention (see McCafferty, J., *et al.*, (1990), *Nature*
10 348, 552-554; Marks, J. *et al.*, (1992) *Biotechnology* 10, 779-783).

Ligands

The polypeptides of the invention may also be used to search for interacting ligands. Methods for doing this include the screening of a library of compounds (see Coligan *et al.*, *Current Protocols in Immunology* 1(2); Chapter 5 (1991), isolating the ligands from cells, isolating the ligands from a cell-free preparation
15 or natural product mixtures. Ligands to the polypeptide may activate (agonise) or inhibit (antagonise) its activity. Alternatively, compounds may affect the levels of the polypeptide present in the cell, including affecting gene expression, mRNA stability and the degree of post-translational modification of the encoded protein. The invention thus embraces methods for the identification of a compound that is effective in the treatment and/or diagnosis of disease, comprising contacting a polypeptide, a nucleic acid
20 molecule or host cell according to any one of the embodiments of the invention described herein with one or more compounds suspected of possessing binding affinity for said polypeptide or nucleic acid molecule, and selecting a compound that binds specifically to said nucleic acid molecule or polypeptide, or that affects the level of gene expression, mRNA stability or the degree of post-translational modification of the encoded protein.

25 Ligands to the polypeptide form a further aspect of the invention, as discussed in more detail above. Preferred "antagonist" ligands include those that bind to the polypeptide of this invention and strongly inhibit any activity of the polypeptide. Preferred "agonist" ligands include those that bind to the polypeptide and strongly induce activity of the polypeptide of this invention or increases substantially the level of the polypeptide in the cell. As defined above, the term "agonist" is meant to include any
30 polypeptide, peptide, synthetic molecule or organic molecule that functions as an activator, by increasing the effective biological activity of a polypeptide, for example, by increasing gene expression or enzymatic activity. The term "antagonist" is meant to include any polypeptide, peptide, synthetic molecule or

organic molecule that functions as an inhibitor, by decreasing the effective biological activity of the gene product, for example, by inhibiting gene expression of an enzyme or a pharmacological receptor.

Ligands to a polypeptide according to the invention may come in various forms, including natural or modified substrates, enzymes, receptors, small organic molecules such as small natural or synthetic
 5 organic molecules of up to 2000Da, preferably 800Da or less, peptidomimetics, inorganic molecules, peptides, polypeptides, antibodies, structural or functional mimetics of the aforementioned.

B. Nucleic acid molecules

Preferred nucleic acid molecules of the invention are those which encode the polypeptide sequences recited in any one of SEQ ID Nos. 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39,
 10 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 63, 67, 69, 73, 75, 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129, 131, 133, 137, 139, 141, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 and 209. Examples of such nucleic acid molecules include those listed in SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74,
 15 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216, homologous nucleic acids and nucleic acids that are complementary to these nucleic acid molecules. Nucleic acid molecules of this aspect of the invention may be used in numerous methods and applications, as described generally herein. A nucleic acid molecule preferably
 20 comprises of at least n consecutive nucleotides from any one of the sequences disclosed in SEQ ID Nos. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192,
 25 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216, where n is 10 or more. A nucleic acid molecule of the invention also includes sequences that are complementary to the nucleic acid molecule described above (for example, for antisense or probing purposes).

A nucleic acid molecule according to this aspect of the invention may be in the form of RNA, such as mRNA, DNA, such as cDNA, synthetic DNA or genomic DNA. The nucleic acid molecule may be
 30 double-stranded or single-stranded. The single-stranded form may be the coding (sense) strand or the non-coding (antisense) strand. A nucleic acid molecule may also comprise an analogue of DNA or RNA, including, but not limited to modifications made to the backbone of the molecule, such as, for example, a peptide nucleic acid (PNA). The term "PNA" as used herein, refers to an antisense molecule that

comprises an oligonucleotide of at least five nucleotides in length linked to a peptide backbone of amino acid residues, preferably ending in lysine. The terminal lysine confers solubility to the composition. PNAs may be pegylated to extend their lifespan in a cell, where they preferentially bind complementary single-stranded DNA and RNA and stop transcript elongation (Nielsen, P.E. *et al.* (1993) *Anticancer Drug Des.* 8:53-63).

A nucleic acid molecule according to this aspect of the invention can be isolated by cloning, purification or separation of the molecule directly from a particular organism, or from a library, such as a genomic or cDNA library. The molecule may also be synthesised, for example, using chemical synthetic techniques such as solid phase phosphoramidite chemical synthesis. RNA may be synthesized *in vitro* or *in vivo* by transcription of the relevant DNA molecule.

Due to the degeneracy of the genetic code, differing nucleic acid sequences may encode the same polypeptide (or mature polypeptide). Thus, nucleic acid molecules included in this aspect of the invention include any molecule comprising a variant of the sequence explicitly recited. Such variants may include variant nucleic acid molecules that code for the same polypeptide (or mature polypeptide) as that explicitly identified, that code for a fragment of the polypeptide, that code for a functional equivalent of the polypeptide or that code for a fragment of the functional equivalent of the polypeptide. Also included in this aspect of the invention, are variant nucleic acid molecules that are derived from nucleotide substitutions, deletions, rearrangements or insertions or multiple combinations of the aforementioned. Such molecules may be naturally occurring variants, such as allelic variants, non-naturally occurring variants such as those created by chemical mutagenesis, or variants isolated from a species, cell or organism type other than the type from which the sequence explicitly identified originated. Variant nucleic acid molecules may differ from the nucleic acid molecule explicitly recited in a coding region, non-coding region or both these regions.

Nucleic acid molecules may also include additional nucleic acid sequence to that explicitly recited, for example, at the 5' or 3' end of the molecule. Such additional nucleic acids may encode for a polypeptide with added functionality compared with the original polypeptide whose sequence is explicitly identified herein. An example of this would be an addition of a sequence that is heterologous to the original nucleic acid sequence, to encode a fusion protein. Such a fusion protein may be of use in aiding purification procedures or enabling techniques to be carried out where fusion proteins are required (such as in the yeast two hybrid system). Additional sequences may also include leader or secretory sequences such as those coding for pro-, pre- or prepro- polypeptide sequences. These additional sequences may also include non-coding sequences that are transcribed but not translated including ribosome binding sites and termination signals.

A nucleic acid molecule of the invention may include molecules that are at least 70% identical over their entire length to a nucleic acid molecule as explicitly identified herein in SEQ ID Nos.: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216. Preferably, a nucleic acid molecule according to this aspect of the invention comprises a region that is at least 80% identical over its entire length to a nucleic acid molecule as explicitly identified herein in these SEQ ID Nos., preferably at least 90%, more preferably at least 95% and most preferably at least 98% or 99% identical. Further preferred embodiments include nucleic acid molecules that encode polypeptides that retain substantially the same biological function or activity as the polypeptide explicitly identified herein. The terms "homology" and "identity" should be given the meanings described in detail above with respect to polypeptide analysis. Preferably, nucleotide homology and identity are assessed using the blastn program available at <http://www.ncbi.nlm.nih.gov>.

The nucleic acid molecules of the invention can also be engineered using methods generally known in the art. These methods include but are not limited to DNA shuffling; random or non-random fragmentation (by restriction enzymes or shearing methods) and reassembly of fragments; insertions, deletions, substitutions and rearrangements of sequences by site-directed mutagenesis (for example, by PCR). These alterations may be for a number of reasons including for ease of cloning (such as introduction of new restriction sites), altering of glycosylation patterns, changing of codon preferences, splice variants changing the processing, and/or expression of the gene product (the polypeptide) in general or creating fusion proteins (see above).

Hybridisation

Nucleic acid molecules of the invention may also include antisense molecules that are partially complementary to a nucleic acid molecule as explicitly identified herein in SEQ ID Nos.: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216, and which therefore will hybridise to the encoding nucleic acid molecules. These antisense molecules, including oligonucleotides, can be designed to recognise, specifically bind to and prevent transcription of a target nucleic acid encoding a polypeptide of the invention, as will be known by those of ordinary skill in the art (see Cohen, J.S., Trends in Pharm. Sci., 10, 435 (1989), Okano, J. Neurochem. 56, 560 (1991); O'Connor, J. Neurochem 56, 560 (1991); Lee

et al., Nucleic Acids Res 6, 3073 (1979); Cooney *et al.*, Science 241, 456 (1988); Dervan *et al.*, Science 251, 1360 (1991).

The term "hybridisation" used herein refers to any process by which a strand of nucleic acid binds with a complementary strand of nucleic acid by hydrogen bonding, typically forming Watson-Crick base pairs.

- 5 As carried out *in vitro*, one of the nucleic acid populations is usually immobilised to a surface, whilst the other population is free. The two molecule types are then placed together under conditions conducive to binding.

- The phrase "stringency of hybridisation" refers to the percentage of complementarity that is needed for duplex formation. "Stringency" thus refers to the conditions in a hybridization reaction that favour the association of very similar molecules over association of molecules that differ. Conditions can therefore exist that allow not only nucleic acid strands with 99-100% complementarity to hybridise, but sequences with lower complementarity (for example, 50%) to also hybridise. High stringency hybridisation conditions are defined herein as overnight incubation at 42°C in a solution comprising 50% formamide, 5XSSC (150mM NaCl, 15mM trisodium citrate), 50mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulphate, and 20 microgram/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1X SSC at approximately 65°C. Low stringency conditions involve the hybridisation reaction being carried out at 35°C (see Sambrook *et al.* [supra]). Preferably, the conditions used for hybridization are those of high stringency.
- 10
15

- Some *trans*- and *cis*-acting factors that may affect the binding of two complementary strands include strand length, base composition (GC pairs have an extra hydrogen bond and are thus require more energy to separate than AT pairs) and the chemical environment. The presence of monovalent cations (such as Na⁺) stabilises duplex formation whereas chemical denaturants such as formamide and urea destabilise the duplex by disruption of the hydrogen bonds. Use of compounds such as polyethylene glycol (PEG) can increase reassociation speeds by increasing overall DNA concentration in aqueous solution by abstracting water molecules. Denhardt's reagent or BLOTTO are chemical agents often added to block non-specific attachment of the liquid phase to the solid support. Increasing the temperature will also increase the stringency of hybridisation, as will increasing the stringency of the washing conditions following hybridisation (Sambrook *et al.* [supra]).
- 20
25

- Numerous techniques exist for effecting hybridisation of nucleic acid molecules. Such techniques usually involve one of the nucleic acid populations being labelled. Labelling methods include, but are not limited to radiolabelling, fluorescence labelling, chemiluminescent or chromogenic labelling or chemically coupling a modified reporter molecule to a nucleotide precursor such as the biotin-streptavidin system.
- 30

This can be done by oligolabelling, nick-translation, end-labelling or PCR amplification using a labelled polynucleotide. Labelling of RNA molecules can be achieved by cloning the sequences encoding the polypeptide of the invention into a vector specifically for this purpose. Such vectors are known in the art and may be used to synthesise RNA probes *in vitro* by the addition of an appropriate RNA polymerase
5 such as T7, T3 or SP6 and labelled nucleotides.

Various kits are commercially available that allow the labelling of molecules. Examples include those made by Pharmacia & Upjohn (Kalamazoo, MI); Promega (Madison WI); and the U.S. Biochemical Corp. (Cleveland, OH). Hybridisation assays include, but are not limited to dot-blots, Southern blotting, Northern blotting, chromosome *in situ* hybridisation (for example, FISH [fluorescence *in situ*
10 hybridisation]), tissue *in situ* hybridisation, colony blots, plaque lifts, gridded clone hybridisation assays, DNA microarrays and oligonucleotide microarrays. These hybridisation methods and others, may be used by a skilled artisan to isolate copies of genomic DNA, cDNA, or RNA encoding homologous or orthologous proteins from other species.

The invention therefore also embodies a process for detecting a nucleic acid molecule according to the
15 invention, comprising the steps of: (a) contacting a nucleic probe with a biological sample under hybridising conditions to form duplexes; and (b) detecting any such duplexes that are formed. The term "probe" as used herein refers to a nucleic acid molecule in a hybridisation reaction whose molecular identity is known and is designed specifically to identify nucleic acids encoding homologous genes in other species. Usually, the probe population is the labelled population, but this is not always the case, as
20 for example, in a reverse hybridisation assay.

One example of a use of a probe is to find nucleic acid molecules with an equivalent function to those that are explicitly identified herein, or to identify additional family members in the same or other species. This can be done by probing libraries, such as genomic or cDNA libraries, derived from a source of interest, such as a human, a non-human animal, other eukaryote species, a plant, a prokaryotic species or a virus.
25 The probe may be natural or artificially designed using methods recognised in the art (for example, Ausubel *et al.*, [*supra*]). A nucleic acid probe will preferably possess greater than 15, more preferably greater than 30 and most preferably greater than 50 contiguous bases complementary to a nucleic acid molecule explicitly identified herein.

In many cases, isolated DNA from cDNA libraries will be incomplete in the region encoding the
30 polypeptide, normally at the 5' end. Methods available for subsequently obtaining full-length cDNA sequence include RACE (rapid amplification of cDNA ends) as described by Frohman *et al.*, (Proc. Natl. Acad. Sci. USA 85, 8998-9002 (1988)), and restriction-site PCR, which uses universal primers to retrieve

- unknown nucleic acid sequence adjacent to a known locus (Sarkar, G. (1993) PCR Methods Applic., 2:318-322). "Inverse PCR" may also be used to amplify or to extend sequences using divergent primers based on a known region (Triglia, T. *et al.*, (1988) Nucleic Acids Res. 16:8186). Another method which may be used is "capture PCR", which involves PCR amplification of DNA fragments adjacent to a known
- 5 sequence in human and yeast artificial chromosome DNA (Lagerstrom, M. *et al.*, (1991) PCR Methods Applic., 1:111-119). Another method which may be used to retrieve unknown sequences is that of Parker, J.D. *et al.*, (1991); Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and libraries, such as the PromoterFinderTM library (Clontech, Palo Alto, CA) to walk genomic DNA. This latter process avoids the need to screen libraries and is useful in finding intron/exon junctions.
- 10 When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Also, random-primed libraries are preferable, in that they will contain more sequences that contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.
- 15 In one embodiment of the invention, a nucleic acid molecule according to the invention may be used for chromosome localisation. In this technique, a nucleic acid molecule is specifically targeted to, and can hybridise with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes is an important step in the confirmatory correlation of those sequences with the gene-associated disease. Once a sequence has been mapped to a precise chromosomal location, the
- 20 physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins University Welch Medical Library). The relationships between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes). This provides valuable information to investigators searching for disease
- 25 genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localised by genetic linkage to a particular genomic region, any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleic acid molecule may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier, or affected individuals.
- 30 Nucleic acid molecules of the present invention are also valuable for tissue localisation. Such techniques facilitate the determination of expression patterns of the polypeptide in tissues by detection of the mRNAs that encode them. These techniques include *in situ* hybridisation techniques and nucleotide amplification techniques, such as PCR. Results from these studies provide an indication of the normal functions of the

polypeptide in the organism, as well as highlighting the involvement of a particular gene in a disease state or abnormal physiological condition.

In addition, comparative studies of the normal expression pattern of mRNAs with that of mRNAs encoded by a mutant gene provide valuable insights into the role of mutant polypeptides in disease. Such
5 inappropriate expression may be of a temporal, spatial or quantitative nature.

Vectors

The nucleic acid molecules of the present invention may be incorporated into vectors for cloning (for example, pBluescript made by Stratagene) or expression purposes. Vectors containing a nucleic acid molecule explicitly identified herein (or a variant thereof) form another aspect of this invention. The
10 nucleic acid molecule may be inserted into an appropriate vector by any variety of well known techniques such as those described in Sambrook *et al.* [supra]. Generally, the encoding gene can be placed under the control of a control element such as a promoter, ribosome binding site or operator, so that the DNA sequence encoding the desired polypeptide is transcribed into RNA in the transformed host cell.

Vectors may be derived from various sources including, but not limited to bacterial plasmids, bacteriophage, transposons, yeast episomes, insertion elements, yeast chromosomal elements, viruses for
15 example, baculoviruses and SV40 (simian virus), vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses, lentiviruses and retroviruses, or combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, including cosmids and phagemids. Human, bacterial and yeast artificial chromosomes (HACs, BACs and YACs respectively) may also be employed to deliver
20 larger fragments of DNA than can be contained and expressed in a plasmid.

Examples of retroviruses include but are not limited to: murine leukaemia virus (MLV), human immunodeficiency virus (HIV), equine infectious anaemia virus (EIAV), mouse mammary tumour virus (MMTV), Rous sarcoma virus (RSV), Fujinami sarcoma virus (FuSV), Moloney murine leukaemia virus (Mo-MLV), FBR murine osteosarcoma virus (FBR MSV), Moloney murine sarcoma virus (Mo-MSV),
25 Abelson murine leukaemia virus (A-MLV), Avian myelocytomatosis virus-29 (MC29), and Avian erythroblastosis virus (AEV). A detailed list of retroviruses may be found in Coffin *et al.* ("Retroviruses" 1997 Cold Spring Harbour Laboratory Press Eds: JM Coffin, SM Hughes, HE Varmus pp 758-763).

Lentiviruses can be divided into primate and non-primate groups. Examples of primate lentiviruses include but are not limited to: the human immunodeficiency virus (HIV), the causative agent of human
30 auto-immunodeficiency syndrome (AIDS), and the simian immunodeficiency virus (SIV). The non-primate lentiviral group includes the prototype "slow virus" visna/maedi virus (VMV), as well as the related caprine arthritis-encephalitis virus (CAEV), equine infectious anaemia virus (EIAV) and the more recently described feline immunodeficiency virus (FIV) and bovine immunodeficiency virus (BIV).

A distinction between the lentivirus family and other types of retroviruses is that lentiviruses have the capability to infect both dividing and non-dividing cells (Lewis et al 1992 EMBO. J 11: 3053-3058; Lewis and Emerman 1994 J. Virol. 68: 510-516). In contrast, other retroviruses - such as MLV - are unable to infect non-dividing cells such as those that make up, for example, muscle, brain, lung and liver
5 tissue.

A vector may be configured as a split-intron vector. A split intron vector is described in PCT patent applications WO 99/15683 and WO 99/15684.

If the features of adenoviruses are combined with the genetic stability of retroviruses/lentiviruses then essentially the adenovirus can be used to transduce target cells to become transient retroviral producer
10 cells that could stably infect neighbouring cells. Such retroviral producer cells engineered to express an antigen of the present invention can be implanted in organisms such as animals or humans for use in the treatment of angiogenesis and/or cancer.

Poxvirus vectors are also suitable for use in accordance with the present invention. Pox viruses are engineered for recombinant gene expression and for the use as recombinant live vaccines. This entails the
15 use of recombinant techniques to introduce nucleic acids encoding foreign antigens into the genome of the pox virus. If the nucleic acid is integrated at a site in the viral DNA which is non-essential for the life cycle of the virus, it is possible for the newly produced recombinant pox virus to be infectious, that is to say to infect foreign cells and thus to express the integrated DNA sequence. The recombinant pox virus prepared in this way can be used as live vaccines for the prophylaxis and/or treatment of pathologic and
20 infectious disease.

For vaccine delivery, preferred vectors are vaccinia virus vectors such as MVA or NYVAC. Most preferred is the vaccinia strain modified virus ankara (MVA) or a strain derived therefrom. Alternatives to vaccinia vectors include avipox vectors such as fowlpox or canarypox known as ALVAC and strains derived therefrom which can infect and express recombinant proteins in human cells but are unable to
25 replicate.

Bacterial vectors may be also used, such as salmonella, listeria and mycobacteria.

Vectors containing the relevant nucleotide sequence may enter the host cell by a variety of methods well known in the art and described in many standard laboratory manuals (such as Sambrook *et al.*, [supra], Ausubel *et al.*, [supra], Davis *et al.*, Basic Methods in Molecular Biology (1986)). Methods include
30 calcium phosphate transfection, cationic lipid-mediated transfection, DEAE-dextran mediated transfection, electroporation, microinjection, scrape loading, transduction, and ballistic introduction or

infection.

Host cells

The choice of host cells is often dependent on the vector type used as a carrier for the nucleic acid molecule of the present invention. Bacteria and other microorganisms are particularly suitable hosts for plasmids, cosmids and expression vectors generally (for example, vectors derived from the pBR322 plasmid), yeast are suitable hosts for yeast expression vectors, insect cell systems are suitable host for virus expression vectors (for example, baculovirus) and plant cells are suitable hosts for vectors such as the cauliflower mosaic virus (CaMV) and tobacco mosaic virus (TMV). Other expression systems include using animal cells (for example, with the LentiVectors™, Oxford BioMedica) as a host cell or even using cell-free translating systems. Some vectors, such as "shuttle vectors" may be maintained in a variety of host cells. An example of such a vector would be pEG 202 and other yeast two-hybrid vectors which can be maintained in both yeast and bacterial cells (see Ausubel *et al.*, [supra] and Gyuris, J., Cell, 75, 791-803).

Examples of suitable bacterial hosts include *Streptococci*, *Staphylococci*, *Escherichia coli*, *Streptomyces* and *Bacillus subtilis* cells. Yeast and fungal hosts include *Saccharomyces cerevisiae* and *Aspergillus* cells. Mammalian cell hosts include many immortalised cell lines available from the American Type Culture Collection (ATCC) such as CHO (Chinese Hamster Ovary) cells, HeLa cells, BHK (baby hamster kidney) cells, monkey kidney cells, C127, 3T3, BHK, HEK 293, Bowes melanoma and human hepatocellular carcinoma (for example, Hep G2) cells. Insect host cells that are used for baculovirus expression include *Drosophila* S2 and *Spodoptera* Sf9 cells. Plant host cells include most plants from which protoplasts be isolated and cultured to give whole regenerated plants. Practically, all plants can be regenerated from cultured cells or tissues, including but not limited to all major species of sugar cane, sugar beet, cotton, fruit and other trees, legumes and vegetables.

Expression systems

Also included in present invention are expression vectors that comprise a nucleic acid molecule as described above. Expression vectors and host cells are preferably chosen to give long term, high yield production and stable expression of the recombinant polypeptide and its variants.

Expression of a polypeptide can be effected by cloning an encoding nucleic acid molecule into a suitable expression vector and inserting this vector into a suitable host cell. The positioning and orientation of the nucleic acid molecule insert with respect to the regulatory sequences of the vector is important to ensure that the coding sequence is properly transcribed and translated. Alternatively, control and other regulatory

sequences may be ligated onto the nucleic acid molecule of this invention prior to its insertion into the expression vector. In both cases, the sequence of the nucleic acid molecule may have to be adjusted in order to effect correct transcription and translation (for example, addition of nucleotides may be necessary to obtain the correct reading frame for translation of the polypeptide from its encoding nucleic acid molecule).

A nucleic acid molecule of the invention may comprise control sequences that encode signal peptides or leader sequences. These sequences may be useful in directing the translated polypeptide to a variety of locations within or outside the host cell, such as to the lumen of the endoplasmic reticulum, to the nucleus, to the periplasmic space, or into the extracellular environment. Such signals may be endogenous to the nucleic acid molecules of the invention, or may be a heterologous sequence. These leader or control sequences may be removed by the host during post-translational processing.

A nucleic acid molecule of the present invention may also comprise one or more regulatory sequences that allow for regulation of the expression of polypeptide relative to the growth of the host cell. Alternatively, these regulatory signals may be due to a heterologous sequence from the vector. Stimuli that these sequences respond to include those of a physical or chemical nature such as the presence or absence of regulatory compounds, changing temperatures or metabolic conditions. Regulatory sequences as described herein, are non-translated regions of sequence such as enhancers, promoters and the 5' and 3' untranslated regions of genes. Regulatory sequences interact with host cellular proteins that carry out translation and transcription. These regulatory sequences may vary in strength and specificity. Examples of regulatory sequences include those of constitutive and inducible promoters. In bacterial systems, an example of an inducible promoter is the hybrid *lacZ* promoter of the Bluescript phagemid (Stratagene, LaJolla, CA) or pSport1™ plasmid (Gibco BRL). The baculovirus polyhedrin promoter may be used in insect cells.

An example of a preferred expression system is the lentivirus expression system, for example, as described in International patent application WO98/17815.

Detection of uptake of vectors by the host organism

Various methods are known in the art to detect the uptake of a nucleic acid or vector molecule by a host cell and/or the subsequent successful expression of the encoded polypeptide (see for example Sambrook *et al.*, [supra]).

Vectors frequently have marker genes that can be easily assayed. Thus, vector uptake by a host cell can be readily detected by testing for the relevant phenotype. Markers include, but are not limited to those coding for antibiotic resistance, herbicide resistance or nutritional requirements. The gene encoding

dihydrofolate reductase (DHFR) for example, confers resistance to methotrexate (Wigler, M. *et al.* (1980) PNAS 77:3567-70) and the gene *npt* confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin, F. *et al.* (1981) J. Mol. Biol. 150:1-14). Additional selectable genes have been described, examples of which will be clear to those of skill in the art.

- 5 Markers however, only indicate that a vector has been taken up by a host cell but does not distinguish between vectors that contain the desired nucleic acid molecule and those that do not. One method of detecting for the said nucleic acid molecule is to insert the relevant sequence at a position that will disrupt the transcription and translation of a marker gene. These cells can then be identified by the absence of a marker gene phenotype. Alternatively, a marker gene can be placed in tandem with a sequence encoding a
- 10 polypeptide of the invention under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

More direct and definitive methods to detect the presence of the nucleic acid molecule of the present invention include DNA-DNA or DNA-RNA hybridisation with a probe comprising the relevant antisense molecule, as described above. More direct methods to detect polypeptide expression include protein

15 bioassays for example, fluorescence activated cell sorting (FACS), immunoassay techniques such as ELISA or radioimmunoassays.

Alternative methods for detecting or quantitating the presence of the nucleic acid molecule or polypeptide of this invention include membrane, solution or chip-based technologies (see Hampton, R. *et al.*, (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN) and Maddox, D.E. *et al.*, (1983) J.

20 Exp. Med, 158, 1211-1216).

Transgenic animals

In another embodiment of this invention, a nucleic acid molecule according to the invention may be used to create a transgenic animal, most commonly a rodent. The modification of the animal's genome may either be done locally, by modification of somatic cells or by germ line therapy to incorporate inheritable

25 modifications. Such transgenic animals may be particularly useful in the generation of animal models for drug molecules effective as modulators of the polypeptides of the present invention.

Polypeptide purification

A polypeptide according to the invention may be recovered and purified from recombinant cell cultures by methods including, but not limited to cell lysis techniques, ammonium sulphate precipitation, ethanol

30 precipitation, acid extraction, anion or cation chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and

lectin chromatography, high performance liquid chromatography (HPLC) or fast performance liquid chromatography (FPLC). The polypeptide may need refolding after purification or isolation and many well known techniques are available that will help regenerate an active polypeptide conformation.

Many expression vectors are commercially available that aid purification of the relevant polypeptide.

- 5 These include vectors that join the sequence encoding the polypeptide to another expressed sequence creating a fused protein that is easier to purify. Ways in which these fused parts can facilitate purification of the polypeptide of this invention include fusions that can increase the solubility of the polypeptide, joining of metal chelating peptides (for example, histidine-tryptophan modules) that allow for purification with immobilised metals, joining of protein A domains which allow for purification with immobilised
- 10 immunoglobulins and the joining of the domain that is utilised in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, WA). Fusion of the polypeptide of this present invention with a secretion signal polypeptide may also aid purification. This is because the medium into which the fused polypeptide has been secreted can subsequently be used to recover and purify the expressed polypeptide.
- 15 If necessary, these extraneous polypeptides often comprise a cleavable linker sequence which allows the polypeptide to be isolated from the fusion. Cleavable linker sequences between the purification domain and the polypeptide of the invention include those specific for Factor Xa or for enterokinase (Invitrogen, San Diego, CA). One such expression vector provides for expression of a fusion protein containing the polypeptide of the invention fused to several histidine residues preceding a thioredoxin or an enterokinase
- 20 cleavage site. The histidine residues facilitate purification by IMAC (immobilised metal ion affinity chromatography as described in Porath, J. *et al.* (1992), *Prot. Exp. Purif.* 3: 263-281), while the thioredoxin or enterokinase cleavage site provides a means for purifying the polypeptide from the fusion protein. A discussion of vectors that contain fusion proteins is provided in Kroll, D.J. *et al.* (1993; *DNA Cell Biol.* 12:441-453).

25 Assays

Another aspect of this invention includes assays that may be carried out using a polypeptide or nucleic acid molecule according to the invention. Such assays may be for many uses including the development of drug candidates, for diagnostic purposes or for the gathering of information for therapeutics.

- If the polypeptide is to be expressed for use in screening assays, generally it is preferred that it be
- 30 produced at the surface of the host cell in which it is expressed. In this event, the host cells may be harvested prior to use in the screening assay, for example using techniques such as fluorescence activated cell sorting (FACS) or immunoaffinity techniques. If the polypeptide is secreted into the medium, the

medium can be recovered in order to recover and purify the expressed polypeptide. If polypeptide is produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

The polypeptide of the invention can be used to screen libraries of compounds in any of a variety of drug screening techniques. Such compounds may activate (agonise) or inhibit (antagonise) the level of
5 expression of the gene or the activity of the polypeptide of the invention and form a further aspect of the present invention. Examples of suitable compounds are those which are effective to alter the expression of a natural gene which encodes a polypeptide of the invention or to regulate the activity of a polypeptide of the invention.

Agonist or antagonist compounds may be isolated from, for example, cells, cell-free preparations,
10 chemical libraries or natural product mixtures. These agonists or antagonists may be natural or modified substrates, ligands, enzymes, receptors or structural or functional mimetics. For a suitable review of such screening techniques, see Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).

Potential agonists or antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to the polypeptide of the invention and thereby modulate its activity. In this fashion, binding of
15 the polypeptide to normal cellular binding molecules may be potentiated or inhibited, such that the normal biological activity of the polypeptide is enhanced or prevented.

The polypeptide of the invention that is employed in such a screening technique may be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. In general, such screening procedures may involve using appropriate cells or cell membranes that express the polypeptide that are
20 contacted with a test compound to observe binding, or stimulation or inhibition of a functional response. The functional response of the cells contacted with the test compound is then compared with control cells that were not contacted with the test compound. Such an assay may assess whether the test compound results in a signal generated by activation of the polypeptide, using an appropriate detection system. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on
25 activation by the agonist in the presence of the test compound is observed.

Alternatively, simple binding assays may be used, in which the adherence of a test compound to a surface bearing the polypeptide is detected by means of a label directly or indirectly associated with the test compound or in an assay involving competition with a labelled competitor. In another embodiment, competitive drug screening assays may be used, in which neutralising antibodies that are capable of
30 binding the polypeptide specifically compete with a test compound for binding. In this manner, the antibodies can be used to detect the presence of any test compound that possesses specific binding affinity for the polypeptide.

Assays may also be designed to detect the effect of added test compounds on the production of mRNA encoding the polypeptide in cells. For example, an ELISA may be constructed that measures secreted or cell-associated levels of polypeptide using monoclonal or polyclonal antibodies by standard methods known in the art, and this can be used to search for compounds that may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues. The formation of binding complexes between the polypeptide and the compound being tested may then be measured.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the polypeptide of interest (see International patent application WO84/03564). In this method, large numbers of different small test compounds are synthesised on a solid substrate, which may then be reacted with the polypeptide of the invention and washed. One way of immobilising the polypeptide is to use non-neutralising antibodies. Bound polypeptide may then be detected using methods that are well known in the art. Purified polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques.

A polypeptide according to the invention may be used to identify membrane-bound or soluble receptors, through standard receptor binding techniques that are known in the art, such as ligand binding and crosslinking assays in which the polypeptide is labelled with a radioactive isotope, is chemically modified, or is fused to a peptide sequence that facilitates its detection or purification, and incubated with a source of the putative receptor (for example, a composition of cells, cell membranes, cell supernatants, tissue extracts, or bodily fluids). The efficacy of binding may be measured using biophysical techniques such as surface plasmon resonance and spectroscopy. Binding assays may be used for the purification and cloning of the receptor, but may also identify agonists and antagonists of the polypeptide, that compete with the binding of the polypeptide to its receptor. Standard methods for conducting screening assays are well understood in the art.

A typical polypeptide-based assay might involve contacting the appropriate cell(s) or cell membrane(s) expressing the polypeptide with a test compound. In such assays, a polypeptide according to the invention may be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. Any response to the test compound, for example a binding response, a stimulation or inhibition of a functional response may then be compared with a control where the cell(s) or cell membrane(s) was/were not contacted with the test compound.

A binding response could be measured by testing for the adherence of a test compound to a surface bearing a polypeptide according to the invention. The test compound may aid polypeptide detection by being labelled, either directly or indirectly. Alternatively, the polypeptide itself may be labelled, for example, with a radioisotope, by chemical modification or as a fusion with a peptide or polypeptide

sequence that will facilitate polypeptide detection. Alternatively, a binding response may be measured, for example, by performing a competition assay with a labelled competitor or *vice versa*. One example of such a technique is a competitive drug screening assay, where neutralising antibodies that are capable of specifically binding to the polypeptide compete with a test compound for binding. In this manner, the
5 antibodies may be used to detect the presence of any test compound that possesses specific binding affinity for the polypeptide. Alternative binding assay methods are well known in the art and include, but are not limited to, cross-linking assays and filter binding assays. The efficacy of binding may be measured using biophysical techniques including surface plasmon resonance and spectroscopy.

High throughput screening is a type of assay which enables a large number of compounds to be searched
10 for any significant binding activity to the polypeptide of interest (see patent application WO84/03564). This is particularly useful in drug screening. In this scenario, many different small test compounds are synthesised on to a solid substrate. The polypeptide is then introduced to this substrate and the whole apparatus washed. The polypeptide is then immobilised by, for example, using non-neutralising antibodies. Bound polypeptide may then be detected using methods that are well known in the art.
15 Purified polypeptide may also be coated directly onto plates for use in the aforementioned drug screening techniques.

Assay methods that are also included within the terms of the present invention are those that involve the use of the genes and polypeptides of the invention in overexpression or ablation assays. Such assays involve the manipulation of levels of these genes/polypeptides in cells and assessment of the impact of
20 this manipulation event on the physiology of the manipulated cells. For example, such experiments reveal details of signaling and metabolic pathways in which the particular genes/polypeptides are implicated, generate information regarding the identities of polypeptides with which the studied polypeptides interact and provide clues as to methods by which related genes and proteins are regulated.

Another aspect of this invention provides for any screening kits that are based or developed from any of
25 the aforementioned assays.

C. Pharmaceuticals

A further aspect of the invention provides a pharmaceutical composition suitable for modulating hypoxia and/or ischaemia, comprising a therapeutically-effective amount of a polypeptide, a nucleic acid molecule, vector or ligand as described above, in conjunction with a pharmaceutically-acceptable carrier.
30 A composition containing a polypeptide, nucleic acid molecule, ligand or any other compound of this present invention (herein known as X) is considered to be "substantially free of impurities" (herein known as Y) when X makes up more than 85% mass per mass of the total [X+Y] mass. Preferably X comprises

at least 90% of the total X+Y mass. More preferably X comprises at least 95%, 98% and most preferably 99% of the total X+Y mass.

Carriers

Carrier molecules may be genes, polypeptides, antibodies, liposomes or indeed any other agent provided
5 that the carrier does not itself induce toxicity effects or cause the production of antibodies that are harmful to the individual receiving the pharmaceutical composition. Further examples of known carriers include polysaccharides, polylactic acids, polyglycolic acids and inactive virus particles. Carriers may also include pharmaceutically acceptable salts such as mineral acid salts (for example, hydrochlorides, hydrobromides, phosphates, sulphates) or the salts of organic acids (for example, acetates, propionates,
10 malonates, benzoates). Pharmaceutically acceptable carriers may additionally contain liquids such as water, saline, glycerol, ethanol or auxiliary substances such as wetting or emulsifying agents, pH buffering substances and the like. Carriers may enable the pharmaceutical compositions to be formulated into tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions to aid intake by the patient. A thorough discussion of pharmaceutically acceptable carriers is available in Remington's Pharmaceutical
15 Sciences (Mack Pub. Co., N.J. 1991).

Dosage

The amount of component X in the composition should also be in therapeutically effective amounts. The phrase "therapeutically effective amounts" used herein refers to the amount of agent needed to treat, ameliorate, or prevent (for example, when used as a vaccine) a targeted disease or condition. An effective
20 initial method to determine a "therapeutically effective amount" may be by carrying out cell culture assays (for example, using neoplastic cells) or using animal models (for example, mice, rabbits, dogs or pigs). In addition to determining the appropriate concentration range for X to be therapeutically effective, animal models may also yield other relevant information such as preferable routes of administration that will give maximum effectiveness. Such information may be useful as a basis for patient administration. A
25 "patient" as used in herein refers to the subject who is receiving treatment by administration of X. Preferably, the patient is human, but the term may also include animals.

The therapeutically-effective dosage will generally be dependent on the patient's status at the time of administration. Factors that may be taken into consideration when determining dosage include the severity of the disease state in the patient, the general health of the patient, the age, weight, gender, diet, time and
30 frequency of administration, drug combinations, reaction sensitivities and the patient's tolerance or response to the therapy. The precise amount can be determined by routine experimentation but may ultimately lie with the judgement of the clinician. Generally, an effective dose will be from 0.01 mg/kg (mass of drug compared to mass of patient) to 50 mg/kg, preferably 0.05 mg/kg to 10 mg/kg.

Compositions may be administered individually to a patient or may be administered in combination with other agents, drugs or hormones.

Routes of administration

Uptake of a pharmaceutical composition of the invention by a patient may be initiated by a variety of methods including, but not limited to enteral, intra-arterial, intrathecal, intramedullary, intramuscular, 5 intranasal, intraperitoneal, intravaginal, intravenous, intraventricular, oral, rectal (for example, in the form of suppositories), subcutaneous, sublingual, transcutaneous applications (for example, see WO98/20734) or transdermal means.

Gene guns or hyposprays may also be used to administer the pharmaceutical compositions of the invention. Typically, the therapeutic compositions may be prepared as injectables, either as liquid 10 solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. Direct delivery of the compositions can generally be accomplished by injection, subcutaneously, intraperitoneally, intravenously or intramuscularly, or delivered to the interstitial space of a tissue. The compositions can also be administered into a lesion. Dosage treatment 15 may be a single dose schedule or a multiple dose schedule.

Inhibition of excessive activity

If a particular disease state is partially or completely caused by an inappropriate excess in the activity of a polypeptide according to the invention, several approaches are available for inhibiting this activity.

One approach comprises administering to a patient an inhibitor compound (antagonist) along with a 20 pharmaceutically acceptable carrier in an amount effective to inhibit the function of the polypeptide, such as by blocking the binding of a ligand, substrate, enzyme, receptor, or by inhibiting a second signal, and thereby alleviating the abnormal condition. Such an antagonist molecule may, for example, be an antibody. Most preferably, such antibodies are chimeric and/or humanised to minimise their immunogenicity, as previously described.

25 In another approach, soluble forms of the polypeptide that retain binding affinity for the ligand, substrate, enzyme, receptor, in question, may be administered to the patient to compete with the biological activity of the endogenous polypeptide. Typically, the polypeptide may be administered in the form of a fragment that retains a portion that is relevant for the desired biological activity.

In an alternative approach, expression of the gene encoding the polypeptide can be inhibited using 30 expression blocking techniques, such as by using antisense nucleic acid molecules (as described above), either internally generated or separately administered. Modifications of gene expression may be effected by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5' or

- regulatory regions (signal sequence, promoters, enhancers and introns) of the gene encoding the polypeptide. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances
- 5 using triplex DNA have been described in the literature (Gee, J.E. *et al.* (1994) In: Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY). The complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes. Such oligonucleotides may be administered or may be generated *in situ* from expression *in vivo*.
- 10 Gene silencing approaches may also be undertaken to down-regulate endogenous expression of a gene. RNA interference (RNAi) (Elbashir, SM *et al.*, Nature 2001, 411, 494-498) is one method of sequence specific post-transcriptional gene silencing that may be employed. Short dsRNA oligonucleotides are synthesised *in vitro* and introduced into a cell. The sequence specific binding of these dsRNA oligonucleotides triggers the degradation of target mRNA, reducing or ablating target protein expression.
- 15 In addition, expression of a polypeptide according to the invention may be prevented by using a ribozyme specific to the encoding mRNA sequence for the polypeptide. Ribozymes are catalytically active RNAs that can be natural or synthetic (see for example Usman, N, *et al.*, Curr. Opin. Struct. Biol (1996) 6(4), 527-33). Synthetic ribozymes can be designed to specifically cleave mRNAs at selected positions thereby preventing translation of the mRNAs into functional polypeptide. Ribozymes may be synthesised with a
- 20 natural ribose phosphate backbone and natural bases, as normally found in RNA molecules. Alternatively the ribozymes may be synthesised with non-natural backbones, for example, 2'-O-methyl RNA, to provide protection from ribonuclease degradation and may contain modified bases.
- Efficacy of the gene silencing approaches assessed above may be assessed through the measurement of polypeptide expression (for example, by Western blotting), and at the RNA level using TaqMan-based
- 25 methodologies.
- RNA molecules may be modified to increase their intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in
- 30 all of these molecules by the inclusion of non-traditional bases such as inosine, queosine and butosine, as well as acetyl-, methyl-, thio- and similarly modified forms of adenine, cytidine, guanine, thymine and uridine that are not as easily recognised by endogenous endonucleases.

Activation of a polypeptide activity

If a particular disease state is partially or completely due to a lowered level of biological activity from a polypeptide according to the invention, various methods may be used. An example of such a method includes administering a therapeutically effective amount of compound that can activate (i.e. an agonist) or cause increased expression of the polypeptide concerned. Administration of such a compound may be via any of the methods described previously.

Gene Therapy

Another aspect of the present invention provides for gene therapy methods involving nucleic acid molecules identified herein. Gene therapy may be used to affect the endogenous production of the polypeptide of the present invention by relevant cells in a patient. For example, gene therapy can be used permanently to treat the inappropriate production of a polypeptide by replacing a defective gene with the corrected therapeutic gene.

Treatment may be effected either *in vivo* or *ex vivo*. *Ex vivo* gene therapy generally involves the isolation and purification of the patient's cells, introduction of the therapeutic gene into the cells and finally, the introduction of the genetically-altered cells back into the patient. *In vivo* gene therapy does not require the isolation and purification of patient cells prior to the introduction of the therapeutic gene into the patient. Instead, the therapeutic gene can be packaged for delivery into the host. Gene delivery vehicles for *in vivo* gene therapy include, but are not limited to, non-viral vehicles such as liposomes, replication-competent and replication-deficient viruses (for example, adenovirus as described by Berkner, K.L., in Curr. Top. Microbiol. Immunol., 158, 39-66 (1992)) or adeno-associated virus (AAV) vectors as described by Muzyczka, N., in Curr. Top. Microbiol. Immunol., 158, 97-129 (1992) and U.S. Patent No. 5,252,479. Alternatively, "naked DNA" may be directly injected into the bloodstream or muscle tissue as a form of *in vivo* gene therapy.

One example of a strategy for gene therapy including a nucleic acid molecule of this present invention may be as follows. A nucleic acid molecule encoding a polypeptide of the invention is engineered for expression in a replication-defective or replication-competent vector, such as a retroviral vector. This expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding the polypeptide, such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a patient for engineering cells *in vivo* and expression of the polypeptide *in vivo* (see Chapter 20, Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, (and references cited therein) in Human Molecular Genetics (1996), T Strachan and A P Read, BIOS Scientific Publishers Ltd).

Genetic delivery of antibodies that bind to polypeptides according to the invention may also be effected, for example, as described in International patent application WO98/55607.

Vaccines

- A further embodiment of the present invention provides that the polypeptides or nucleic acid molecules identified may be used in the development of vaccines. Where the aforementioned polypeptide or nucleic acid molecule is a disease-causing agent, vaccine development can involve the raising of antibodies against such agents. Where the aforementioned polypeptide or nucleic acid molecule is that is upregulated, vaccine development can involve the raising of antibodies or T cells against such agents (as described in WO00/29428).
- 10 Vaccines according to the invention may either be prophylactic (i.e. prevents infection) or therapeutic (i.e. treats disease after infection). Such vaccines comprise immunising antigen(s), immunogen(s), polypeptide(s), protein(s) or nucleic acid, usually in combination with pharmaceutically-acceptable carriers as described above. Additionally, these carriers may function as immunostimulating agents ("adjuvants"). Furthermore, the antigen or immunogen may be conjugated to a bacterial toxoid, such as a
- 15 toxoid from diphtheria, tetanus, cholera, *H. pylori*, and other pathogens.

Vaccination processes may involve the use of heterologous vectors eg: prime with MVA and boost with DNA.

- Since polypeptides may be broken down in the stomach, vaccines comprising polypeptides are preferably administered parenterally (for instance, subcutaneous, intramuscular, intravenous, or intradermal
- 20 injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions that may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the recipient, and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents.

- The vaccine formulations of the invention may be presented in unit-dose or multi-dose containers. For
- 25 example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

The technology referred to as jet injection (see, for example, www.powderject.com) may also be useful in the formulation of vaccine compositions.

- 30 In accordance with this aspect of the present invention, polypeptides can be delivered by viral or non-viral techniques. Non-viral delivery systems include but are not limited to DNA transfection methods. Here, transfection includes a process using a non-viral vector to deliver a antigen gene to a target mammalian

cell. Typical transfection methods include electroporation, nucleic acid biolistics, lipid-mediated transfection, compacted nucleic acid-mediated transfection, liposomes, immunoliposomes, lipofectin, cationic agent-mediated, cationic facial amphiphiles (CFAs) (Nature Biotechnology 1996 14; 556), multivalent cations such as spermine, cationic lipids or polylysine, 1, 2-bis (oleoyloxy)-3-
5 (trimethylammonio) propane (DOTAP)-cholesterol complexes (Wolff and Trubetskoy 1998 Nature Biotechnology 16: 421) and combinations thereof.

Viral delivery systems include but are not limited to adenovirus vectors, adeno-associated viral (AAV) vectors, herpes viral vectors, influenza, retroviral vectors, lentiviral vectors or baculoviral vectors, venezuelan equine encephalitis virus (VEE), poxviruses such as: canarypox virus (Taylor et al 1995
10 Vaccine 13:539-549), entomopox virus (Li Y et al 1998 XIIth International Poxvirus Symposium p144. Abstract), penguin pox (Standard et al. J Gen Virol. 1998 79:1637-46) alphavirus, and alphavirus based DNA vectors.

In addition to the use of polypeptide-based vaccines, this aspect of the invention includes the use of genetically-based vaccines, for example, those vaccines that are effective through eliciting the expression
15 of a particular gene (either endogenous or exogenously derived) in a cell, so targeting this cell for destruction by the immune system of the host organism.

A number of suitable methods for vaccination and vaccine delivery systems are described in International patent application WO00/29428.

D. Diagnostics

20 Another aspect of the present invention provides for the use of a nucleic acid molecule identified herein as a diagnostic reagent.

For example, a nucleic acid molecule may be detected or isolated from a patient's tissue and used for diagnostic purposes. "Tissue" as defined herein refers to blood, urine, any matter obtained from a tissue biopsy or any matter obtained from an autopsy. Genomic DNA from the tissue sample may be used
25 directly for detection of a hypoxia-related condition. Alternatively, the DNA may be amplified using methods such as polymerase chain reaction (PCR), the ligase chain reaction (LCR), strand displacement amplification (SDA), or other amplification techniques (see Saiki *et al.*, Nature, 324, 163-166 (1986); Bej, *et al.*, Crit. Rev. Biochem. Molec. Biol., 26, 301-334 (1991); Birkenmeyer *et al.*, J. Virol. Meth., 35, 117-126 (1991) and Brunt, J., Bio/Technology, 8, 291-294 (1990)). Such diagnostics are particularly
30 useful for prenatal and even neonatal testing.

A method of diagnosis of disease using a polynucleotide may comprise assessing the level of expression of the natural gene and comparing the level of encoded polypeptide to a control level measured in a

normal subject that does not suffer from the disease or physiological condition that is being tested. The diagnosis may comprise the following steps:

- 5 a) contacting a sample of tissue from the patient with a nucleic acid probe under stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule of the invention and the probe;
- b) contacting a control sample with said probe under the same conditions used in step a); and
- c) detecting the presence of hybrid complexes in said samples;

wherein detection of differing levels of the hybrid complex in the patient sample compared to levels of the hybrid complex in the control sample is indicative of the dysfunction.

10 A further aspect of the invention comprises a diagnostic method comprising the steps of:

- a) obtaining a tissue sample from a patient being tested for disease;
- b) isolating a nucleic acid molecule according to the invention from said tissue sample; and
- c) diagnosing the patient for disease by detecting the presence of a mutation in the nucleic acid molecule which is associated with disease.

- 15 To aid the detection of nucleic acid molecules in the above-described methods, an amplification step, such as PCR, may be included. An example of this includes detection of deletions or insertions indicative of the dysfunction by a change in the size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridising amplified DNA to labelled RNA of the invention or alternatively, labelled antisense DNA sequences of the invention. Perfectly matched sequences can be
- 20 distinguished from mismatched duplexes by RNase digestion or by assessing differences in melting temperatures. The presence or absence of the mutation in the patient may be detected by contacting DNA with a nucleic acid probe that hybridises to the DNA under stringent conditions to form a hybrid double-stranded molecule, the hybrid double-stranded molecule having an unhybridised portion of the nucleic acid probe strand at any portion corresponding to a mutation associated with disease; and detecting the
- 25 presence or absence of an unhybridised portion of the probe strand as an indication of the presence or absence of a disease-associated mutation in the corresponding portion of the DNA strand.

- Point mutations and other sequence differences between the reference gene and "mutant" genes can be identified by other well-known techniques, such as direct DNA sequencing or single-strand conformational polymorphism, (see Orita *et al.*, *Genomics*, 5, 874-879 (1989)). For example, a
- 30 sequencing primer may be used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabelled nucleotides or by automatic sequencing procedures with fluorescent-tags.

Cloned DNA segments may also be used as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. Further, point mutations and other sequence variations, such as polymorphisms, can be detected as described above, for example, through the use of allele-specific oligonucleotides for PCR amplification of sequences that differ by single nucleotides.

- 5 DNA sequence differences may also be detected by alterations in the electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing (for example, Myers *et al.*, Science (1985) 230:1242). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (see Cotton *et al.*, PNAS, USA (1985) 85: 4397-4401).
- 10 In addition to conventional gel electrophoresis and DNA sequencing, mutations such as microdeletions, aneuploidies, translocations, inversions, can also be detected by *in situ* analysis (see, for example, Keller *et al.*, DNA Probes, 2nd Ed., Stockton Press, New York, N.Y., USA (1993)), that is, DNA or RNA sequences in cells can be analysed for mutations without need for their isolation and/or immobilisation onto a membrane. FISH is presently the most commonly applied method and numerous reviews of FISH
- 15 have appeared (see, for example, Trachuck *et al.*, Science, 250, 559-562 (1990), and Trask *et al.*, Trends, Genet., 7, 149-154 (1991)).

Arrays

- In another embodiment of the invention, an array of oligonucleotide probes comprising a nucleic acid molecule according to the invention can be constructed to conduct efficient screening of genetic variants,
- 20 mutations and polymorphisms. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see for example: M.Chee *et al.*, Science (1996), Vol 274, pp 610-613).

- In one embodiment, the array is prepared and used according to the methods described in WO95/11995 (Chee *et al.*); Lockhart, D. J. *et al.* (1996) Nat. Biotech. 14: 1675-1680); and Schena, M. *et al.* (1996)
- 25 PNAS 93: 10614-10619). Oligonucleotide pairs may range from two to over one million. The oligomers are synthesized at designated areas on a substrate using a light-directed chemical process. The substrate may be paper, nylon or other type of membrane, filter, chip, glass slide or any other suitable solid support. In another aspect, an oligonucleotide may be synthesized on the surface of the substrate by using a chemical coupling procedure and an ink jet application apparatus, as described in PCT application
- 30 WO95/251116 (Baldeschweiler *et al.*). In another aspect, a "gridded" array analogous to a dot (or slot) blot may be used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical bonding procedures. An array, such as those described above, may be produced by hand or by using available devices (slot blot or dot blot apparatus),

materials (any suitable solid support), and machines (including robotic instruments), and may contain 8, 24, 96, 384, 1536 or 6144 oligonucleotides, or any other number between two and over one million which lends itself to the efficient use of commercially-available instrumentation.

Diagnostics using polypeptides or mRNA

- 5 In addition to the methods discussed above, diseases may be diagnosed by methods comprising determining, from a sample derived from a subject, an abnormally decreased or increased level of polypeptide or mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, nucleic acid amplification, for instance PCR, RT-PCR, RNase protection, Northern blotting and other
10 hybridization methods.

- Assay techniques that can be used to determine levels of a polypeptide of the present invention in a sample derived from a host are well-known to those of skill in the art and are discussed in some detail above (including radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays). One example of this aspect of the invention provides a diagnostic method which comprises the
15 steps of: (a) contacting a ligand as described above with a biological sample under conditions suitable for the formation of a ligand-polypeptide complex; and (b) detecting said complex.

- Protocols such as ELISA, RIA, and FACS for measuring polypeptide levels may additionally provide a basis for diagnosing altered or abnormal levels of polypeptide expression. Normal or standard values for polypeptide expression are established by combining body fluids or cell extracts taken from normal
20 mammalian subjects, preferably humans, with antibody to the polypeptide under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, such as by photometric means.

- Antibodies which specifically bind to a polypeptide of the invention may be used for the diagnosis of conditions or diseases characterised by expression of the polypeptide, or in assays to monitor patients
25 being treated with the polypeptides, nucleic acid molecules, ligands and other compounds of the invention. Antibodies useful for diagnostic purposes may be prepared in the same manner as those described above for therapeutics. Diagnostic assays for the polypeptide include methods that utilise the antibody and a label to detect the polypeptide in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labelled by joining them, either
30 covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules known in the art may be used, several of which are described above.

Quantities of polypeptide expressed in subject, control and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the

parameters for diagnosing disease. Diagnostic assays may be used to distinguish between absence, presence, and excess expression of polypeptide and to monitor regulation of polypeptide levels during therapeutic intervention. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials or in monitoring the treatment of an individual patient.

Diagnostic kits

A diagnostic kit of the present invention may comprise:

- (a) a nucleic acid molecule of the present invention;
- (b) a polypeptide of the present invention; or
- 10 (c) a ligand of the present invention.

In one aspect of the invention, a diagnostic kit may comprise a first container containing a nucleic acid probe that hybridises under stringent conditions with a nucleic acid molecule according to the invention; a second container containing primers useful for amplifying the nucleic acid molecule; and instructions for using the probe and primers for facilitating the diagnosis of disease. The kit may further comprise a third container holding an agent for digesting unhybridised RNA.

In an alternative aspect of the invention, a diagnostic kit may comprise an array of nucleic acid molecules, an array of antibody molecules, and/or an array of polypeptide molecules, as discussed in more detail above.

Such kits will be of use in diagnosing a disease or susceptibility to disease, particularly inflammation, oncology, or cardiovascular disease.

Various aspects and embodiments of the present invention will now be described in more detail by way of example, with particular reference to polypeptides regulated differentially under hypoxic conditions as opposed to normoxic conditions. It will be appreciated that modification of detail may be made without departing from the scope of the invention.

25 **Brief description of the Figures**

Figure 1 shows a scatter plot, showing normalised signal intensities in hypoxia versus normoxia, with each dot representing a single gene.

Figure 2: Hypoxia responses amplified by HIF1alpha overexpression. Data shown is the average of 4 repeat experiments. Values represent fold change as compared to untreated cells (condition 1). Error bars represent standard error of the mean.

Figure 3: Hypoxia responses amplified by EPAS1 overexpression. Data shown is the average of 4 repeat experiments. Values represent fold change as compared to untreated cells (condition 1). Error bars represent standard error of the mean.

Figure 4: Hypoxia responses amplified by HIF1alpha / EPAS1 overexpression. Data shown is the average
5 of 4 repeat experiments. Values represent fold change as compared to untreated cells (condition 1). Error bars represent standard error of the mean.

Figure 5 shows genes that are induced by hypoxia to a greater degree in resting macrophages, as compared to activated macrophages. Error bars show the standard deviation from both repeat experiments and multiple exposures from single experiments. These data are not shown in table form. All bars are
10 ratios of mRNA expression in hypoxia/ normoxia. These are calculated separately for resting (light bars) and activated (dark bars) macrophages, and do not illustrate differences resulting from activation in normoxia.

Figure 6 shows genes which are induced by hypoxia to a greater degree in activated macrophages, compared to resting macrophages.

15 Figure 7 shows genes that are repressed by hypoxia to a greater degree in activated macrophages.

For Figures 8, 9a, 9c, 10-32a, 32d and 33-52, mRNA levels, determined from a custom gene array, of particular genes are shown on the Y-axis, expressed as a value as compared to the median expression level of this gene throughout all samples. Eleven primary human cell types as shown on the x-axis were cultured in normoxia (black), or exposed to hypoxia for 6hr (grey) or 18hr (white).

20 Figure 8: Ecotropic viral integration site 2A (Seq ID:475/476).

Figure 9a: Novel PI-3-kinase adapter (Seq ID:79/80); Image clone accession R62339.

Figure 9b: TaqMan Real-time Q-RT-PCR data for Novel PI-3-kinase adapter (Seq ID:79/80); Image clone accession R62339.

Figure 9c: IMAGE clone acc R59598 (Syk).

25 Figure 10: Regulator of G-protein signalling 1 (Seq ID:375/376)

Figure 11: GM2 ganglioside activator protein (Seq ID:389/390)

Figure 12: Hypothetical protein PRO0823 (Seq ID:21/22)

Figure 13: CYP1 (cytochrome P450, subfamily XXVIIIB) (Seq ID:339/340)

Figure 14: Alpha-2-macroglobulin (Seq ID:405/406)

30 Figure 15: Interleukin 1 receptor antagonist (Seq ID:357/358)

Figure 16: SCYA3L (Seq ID:469/470)

Figure 17: CFFM4 (Seq ID:433/434)

Figure 18: Pleckstrin (Seq ID:431/432)

Figure 19: CYP1B1 (SeqID:325/326)

5 Figure 20: CYP1B1 (SeqID:137/138)

Figure 21: Hypothetical protein FLJ13511 (SeqID:163/164)

Figure 22: Hematopoietic Zinc finger protein (SeqID:17/18)

Figure 23: Osteopontin (SeqID:267/268)

Figure 24: Osteopontin (SeqID:267/268)

10 Figure 25: Adipophilin (SeqID:313/314)

Figure 26: Adipophilin (SeqID:313/314)

Figure 27: Adipophilin (SeqID:313/314)

Figure 28: Adipophilin (SeqID:313/314)

Figure 29: Hypothetical protein FLJ22690 (SeqID:205/206)

15 Figure 30: cDNA DKFZp586E1624 (SeqID: 65/66)

Figure 31: EST (SeqID:197/198)

Figure 32a: EGL nine (*C.elegans*) homolog.3 (SeqID:85/86)

Figure 32b: Gene expression profiles in macrophages with and without activation. mRNA levels, determined from a custom gene array, of *c1orf12* are shown on the Y-axis, expressed as a value compared to the mean value of a set of control genes on each array (per-chip normalisation). All cells were human macrophages, cultured either without cytokines or with IL-10 or with the combination of IFN γ and LPS in normoxia and hypoxia.

20

Figure 32c: Gene expression profiles in macrophages with and without activation. mRNA levels, determined from a custom gene array, of *EGLN3* are shown on the Y-axis, expressed as a value compared to the mean value of a set of control genes on each array (per-chip normalisation). All cells were human macrophages, cultured either without cytokines or with IL-10 or with the combination of IFN γ and LPS in normoxia and hypoxia.

25

Figure 32d: *C1orf12* (SeqID: 89.90)

Figure 32e: The effect of EPAS/ HIF overexpression on expression of the gene Clorf12 EGLN genes using a custom gene array. mRNA expression levels of the gene cLORF12 as determined by the custom array, in response to hypoxia and adenoviral over-expression of HIF or EPAS are shown. Experimental conditions are as follows: #1 no adeno / normoxia; #2 empty adeno (low dose)/ normoxia; #3 empty adeno (high dose)/ normoxia; #4 empty adeno (low dose)/ hypoxia; #5 empty adeno (high dose)/ hypoxia; #6 HIF-1 adeno (low dose)/ hypoxia; #7 HIF-1 adeno (high dose)/ hypoxia; #8 EPAS adeno (low dose)/ hypoxia; #9 EPAS adeno (high dose)/ hypoxia. Error bars are the standard error of the mean.

Figure 32f: The effect of EPAS/ HIF overexpression on expression of the gene EGLN3 gene using a custom gene array. mRNA expression levels of the gene EGLN3 as determined by the custom array, in response to hypoxia and adenoviral over-expression of HIF or EPAS are shown. Experimental conditions are as follows: #1 no adeno / normoxia; #2 empty adeno (low dose)/ normoxia; #3 empty adeno (high dose)/ normoxia; #4 empty adeno (low dose)/ hypoxia; #5 empty adeno (high dose)/ hypoxia; #6 HIF-1 adeno (low dose)/ hypoxia; #7 HIF-1 adeno (high dose)/ hypoxia; #8 EPAS adeno (low dose)/ hypoxia; #9 EPAS adeno (high dose)/ hypoxia. Error bars are the standard error of the mean.

Figure 32g: The effect of EPAS/ HIF overexpression on expression of the EGLN3 gene using AffyMetrix Hu95 ver2 GeneChips. mRNA expression levels of the gene in response to hypoxia and adenoviral over-expression of HIF or EPAS are shown. Graphs show the mean of two replicate arrays, with error bars as standard deviation. Above each graph, data values are shown, including the normalised values and raw values (the AffyMetrix average difference parameter) and Present/ Absent flags.

Figure 32h: The effect of EPAS/ HIF overexpression on expression of the clorf12 gene using AffyMetrix Hu95 ver2 GeneChips. mRNA expression levels of the gene in response to hypoxia and adenoviral over-expression of HIF or EPAS are shown. Graphs show the mean of two replicate arrays, with error bars as standard deviation. Above each graph, data values are shown, including the normalised values and raw values (the AffyMetrix average difference parameter) and Present/ Absent flags.

Figure 32i: Flag immunocytochemistry in HEK293T cells

Figure 32j: Human Cardiomyocyte Caspase Activity after 72 hours transduction with EIAV-ELG9-Homolog 3

Figure 33: Novel Metallothionein (SeqID:83/84)

Figure 34: Hypothetical protein hqp0376 (SeqID:337/338)

Figure 35: Metallothionein 2A (SeqID:265/266)

Figure 36: Metallothionein 1G (SeqID:243/244)

Figure 37: Metallothionein 1H (SeqID: 239/240)

Figure 38: Hepcidin antimicrobial peptide (SeqID:141/142)

Figure 39: EST (SeqID: 117/118)

Figure 40: Hypothetical protein FLJ22622 (SeqID:129/130)

Figure 41: TRIP-Br2 (SeqID:31/32)

5 Figure 42: Tumor protein D52 (SeqID:301/302)

Figure 43: Semaphorin 4b (SeqID:91/92/92a)

Figure 44: Dec-1 (SeqID:371/372)

Figure 45: Calgranulin A (SeqID:447/448)

Figure 46: ERO1 (*S. cerevisiae*)-like (SeqID:67/68)

10 Figure 47: Hypothetical protein FLJ20500 (SeqID:25/26)

Figure 48: N-myc downstream regulated (SeqID:229/230)

Figure 49: Decidual protein induced by progesterone (SeqID:387/388)

Figure 50: Integrin, alpha 5 (SeqID:379/380)

Figure 51: Tissue factor (SeqID:225/226)

15 Figure 52: COX-2 (SeqID:237/238)

Figure 53: Genes up-regulated by macrophage activation. Normalised mRNA levels in the 6 experimental conditions (#1 no cytokines/ normoxia, #2 no cytokines/ hypoxia, #3 IL-10/ normoxia, #4 IL-10/ hypoxia, #5 LPS/IFN/ normoxia, #6 LPS/IFN/ hypoxia) are shown as values referenced to the median value of that gene throughout all 6 experimental conditions. Error bars show the standard error of the mean.

20 Figure 54: Genes downregulated by macrophage activation (I)

Figure 55: Genes downregulated by macrophage activation (II)

Figure 56: Genes downregulated by macrophage activation (III)

Figure 57 shows an RNase protection assay for the gene encoding Semaphorin 4b.

Figure 58 shows a Northern blot showing the size of the mRNA and tissue distribution for the

25 Semaphorin 4b gene.

Examples

Summary

Subtracted cDNA libraries were separately prepared for hypoxic macrophages and cardiomyoblasts. This involved harvesting RNA from cells both in normoxia and hypoxia, and preparing cDNA. Subtractive
5 hybridization / suppression PCR was then performed to remove genes from the hypoxic cell cDNA, which are also present in cDNA from normoxic cells. Insert DNA from the libraries was PCR amplified and arrayed onto duplicate membranes. Quantitative hybridizations with pre-library cDNA material (normoxia and hypoxia) were done to identify clones in the libraries that actually contain hypoxia inducible genes. The insert DNA was then sequenced.

- 10 This procedure was done independently for macrophage and cardiomyoblast. The hypoxia inducible genes identified from these different cell types differed widely, with only a minority of these genes being identified from both cell types.

To characterise the differences between the two tissues further, arrays were produced containing all confirmed hypoxia-inducible genes from the macrophage library. Replicate arrays were hybridised with
15 cDNA from normoxic and hypoxic cardiomyoblasts to allow quantitative evaluation of these genes in the cardiomyoblast. This revealed quantitative differences in the hypoxia induced activation these genes in the two cell types.

Example 1a: Comparison of the hypoxic-response between human macrophages and cardiomyoblasts by a subtraction cloning / array screening approach

20 Methods / Results

To isolate human macrophage, monocytes were derived from peripheral blood of healthy human donors. 100ml bags of buffy coat from the Bristol Blood Transfusion Centre were mixed with an equal volume of RPMI1640 medium (Sigma). This was layered on top of 10ml ficol-paque (Pharmacia) in 50ml centrifuge tubes and centrifuged for 25 min at 800 x g. The interphase layer was removed, washed in MACS buffer
25 (phosphate buffered saline pH 7.2, 0.5% bovine serum albumin, 2mM EDTA) and resuspended at 80 microliter per 10⁷ cells. To this 20 microliter CD14 Microbeads (Miltenyi Biotec) were added, and the tube incubated at 4 degrees for 15 min. Following this one wash was performed in MACS buffer at 400 x g and the cells were resuspended in 3 ml MACS buffer and separated on an LS+ MACS Separation Column (Miltenyi Biotec) positioned on a midi-MACS magnet (Miltenyi Biotec). The column was
30 washed with 3 x 3ml MACS buffer. The column was removed from the magnet and cells were eluted in 5 ml MACS buffer using a syringe. Cells were washed in culture medium (AIM V (Sigma) supplemented with 2% human AB serum (Sigma), and resuspended at 2 x 10⁵ cells per ml in the same medium and

placed in large teflon-coated culture bags (Sud-Laborbedarf GmbH, 82131 Gauting, Germany) and transferred to a tissue culture incubator (37 degrees, 5% CO₂) for 7-10 days. During this period monocytes spontaneously differentiate to macrophages. This is confirmed by examining cell morphology using phase contrast microscopy. Cells are removed from the bags by placing at 4 degrees for 30 min and
 5 emptying the contents. The cells are then washed and resuspended in culture medium at 5×10^5 cell/ml and plated out in Primaria 10 cm tissue culture petri dishes (Falcon Becton Dickinson) at 5×10^6 cells per dish. Culture is continued for 16-24hr to allow cell adherence, prior to experimentation involving hypoxia.

As an alternative primary cell type human cardiomyoblast cultures were established. Cells derived from
 10 the ventricular tissue of newborn or foetal hearts were purchased from BioWhittaker (CC-2582). Growth conditions were used to allow maximum expansion of the cells in vitro, by using a medium rich in growth factors. Under such conditions cardiomyoblast-like cells predominate (the developmental precursor of cardiomyocytes). This has been previously described by Goldman and Wurzel (*In Vitro Cell. Dev. Biol.* 28A: 109-119 (1992)) and Goldman *et al.*, (1996, *Exp. Cell. Res.* 228(2): 237-245).

15 For these cultures, cells were seeded at 1×10^6 per T150 flask in human smooth muscle growth medium (TCS CellWorks ZHM-3935) and were expanded in the same medium up to a maximum number of 4 passages. The growth medium is purchased pre-prepared, and includes in the formula, 5% fetal bovine serum, insulin, epidermal growth factor and fibroblast growth factor. Prior to experimentation involving hypoxia, cells were plated onto 10 cm tissue culture petri dishes and allowed to reach confluency.

20 For experimentation with hypoxia, for all cell types, an equal number of identical culture dishes were divided into two separate incubators: One at 37 degrees, 5% CO₂, 95% air (=Normoxia) and the other at 37 degrees, 5% CO₂, 94.9% Nitrogen, 0.1% Oxygen (=Hypoxia). After 6 hours culture under these conditions, the dishes were removed from the incubator, placed on a chilled platform, washed in cold PBS and total RNA was extracted using RNazol B (Tel-Test, Inc; distributed by Biogenesis Ltd) following the
 25 manufacturer's instructions. Polyadenylated mRNA was extracted from the total RNA using a commercial kit following the manufacturer's instructions (Promega; PolyA Tract mRNA isolation System IV).

The hypoxia period of 6 hr was previously determined to be sufficient to allow the induction of known hypoxia-regulated genes, as determined by RNase protection assays. During these preliminary studies it
 30 was noted that macrophages, cardiomyoblasts and an additional control cell type, Jurkat T-cells, showed different patterns of gene induction in response to hypoxia:

Known Hypoxia-inducible gene

level of hypoxia-induced increase in mRNA levels

Macrophage Myoblast T-cell

	phosphoglycerate kinase-1 (PGK)	none	none	high
	vascular endothelial growth factor-A (VEGF)	high	low	high
5	solute carrier family 2, member 1 (Glut-1)	high	low	high

- Separate subtracted cDNA populations were generated from mRNA extracted from hypoxic macrophages and hypoxic cardiomyoblasts, using a combination of two kits, purchased from Clontech Laboratories- SMART PCR cDNA synthesis kit and PCR Select cDNA subtraction kit. The manufacturer's instructions
- 10 were followed for both kits. All diagnostic steps were followed as recommended by the manufacturers. All PCR reactions were done using an Applied Biosystems 9700 with 96-well block, using Applied Biosystems plastics. Driver and tester populations used for subtraction were as below:

subtracted cDNA	tester	driver
Subtracted macrophage	macrophage (hypoxia)	macrophage (normoxia)
Subtracted cardiomyoblast	cardiomyoblast (hypoxia)	cardiomyoblast (normoxia)

- 15 The final subtracted cDNA samples were evaluated by performing RT-PCR using the following primers for human beta actin:

sense: TCACCCACACTGTGCCCATCTACGA
 antisense: CAGCGGAACCGCTCATTGCCAAATGG

- This showed that an additional 5 cycles of PCR were required to achieve similar levels of beta actin
- 20 product from subtracted compared to unsubtracted cDNA, indicating a significant reduction in the representation of a non-regulated gene in the subtracted cDNA. Glyceraldehyde 3-Phosphate dehydrogenase PCR primers, as contained in the kit, were not used.

- The three subtracted cDNA populations were ligated into a plasmid vector (pCRII, Invitrogen) to generate libraries, which were transformed into *E.coli* (INVαF', Invitrogen) and plated out onto agar,
- 25 supplemented with ampicillin and X-Gal, according to standard methods.

Colonies that are white indicate the presence of a recombinant plasmid, and these were picked into individual wells of 96-well plates containing 100 microliters LB-Ampicillin, and given 3-8 hr growth at 37 degrees. In this way, for each library, up to 15 x 96-well plates of clones were generated.

To screen clones for the presence of differentially expressed genes, replicate arrays of plasmid insert DNA were generated on nylon membranes: Firstly, PCR was performed using nested PCR primers 2R and 1, which flank the cDNA insert of each clone (sequence described in the PCR Select kit). The reaction mix also contains 200 μ M d(A,T,C,G)TP, Advantage2 polymerase mix (Clontech Laboratories) and supplied 10x buffer. 40 μ l reactions were set up in 96-well PCR reaction plates and inoculated with 0.5 μ l bacteria from the library plates. 23 cycles of PCR were performed (95 degrees 10 sec; 68 degrees 2 min), and a selection of wells were checked on an agarose gel. In this manner a 96-well plate of insert DNA was generated for each 96-well plate of bacterial clones. Arrays of insert DNA were generated by transferring 4 μ l of each well to 384-well plates (Genetix), and denaturing the DNA by adding 4 μ l 0.4M NaOH and incubating at 37 degrees for 15 minutes. Bromophenol blue was added to the wells to allow visualisation of arraying. A 384-pin replicator (Genetix) was used to spot small volumes of denatured insert DNA onto dry nylon membranes (Hybond N+, Amersham Pharmacia).

By repeating this operation from the same 384-well plate onto several membranes, matched pairs of membranes were produced, suitable for array screening. A fragment of the beta actin gene was spotted at specific positions of the arrays. Following spotting, the membranes were left at room temperature for 2 hr, re-denatured by placing on chromatography paper wetted with 0.3 M NaOH, neutralised by placing on chromatography paper wetted with 0.5 M Tris pH 7.5, dried at room temperature for 2 hr and crosslinked by exposing to 2000 joules UV radiation. Prior to hybridisation, residual salts were removed from the arrays, by washing in hot 0.5% SDS.

Matched pairs of membranes were hybridised with subtracted cDNA samples; from hypoxic and normoxic cells, to determine the abundance of the genes corresponding to each spotted clone in the cDNA samples. Because the cDNA probes were subtracted, large differences in the hybridisation signal for individual spots were apparent, which can be identified by eye. Prior to probe labelling, subtracted cDNA samples were digested with RsaI and run through Qiagen Qiaquick PCR purification columns to remove adapter sequences added during the PCR Select procedure. 25 ng cDNA was labelled with 33 P using a commercial kit following the manufacturer's instructions (Promega, Prime-a-gene kit), and unincorporated label was removed using BioRad Biospin-6 columns following adding 2.5 μ g yeast tRNA carrier.

Pre-hybridisation, hybridisation and washes were performed essentially according to the Research Genetics GeneFilters protocol, but supplementing the hybridisation mixture with 10 μ g of a cocktail of oligonucleotides complementary to the Clontech PCR Select nested PCR primers (equimolar mix of primers 1 and 2R and their reverse complements).

Hybridized arrays were exposed to X-ray film or were exposed to a phosphorimager (Molecular Dynamics, Storm) and clones showing gross differences in the hybridization signals with hypoxic compared to normoxic cDNA probes were identified. This procedure was used to process all clones originally picked from the primary libraries and PCR amplified. The selected clones were grouped
5 together onto a single array (referred to here as a secondary array), and were re-screened with cDNA probes which had not been subtracted, to allow a more quantitative though less sensitive, evaluation of the relative abundance of the genes in hypoxia vs. normoxia.

In this case, probes were ds cDNA generated from the Clontech SMART cDNA synthesis kit (labelled using the Promega Prime-a-gene kit) or were total RNA (labelled according to the Research Genetics
10 GeneFilters protocol), and hybridisations were done according to the Research Genetics GeneFilters protocol.

Hybridization signals were measured using a phosphorimager and were processed with ArrayVision (Imaging Research Inc) software using multiple beta-actin spots to normalise the quantitation and individual spot background correction. At this stage, the inserts of clones showing consistent up-
15 regulation in hypoxia were sequenced using the 2R primer.

The identity of the genes were determined using BLAST at the NCBI (NLM, NIH) against the non-redundant data base collection. Where significant matches to human genes were not made, the human EST database was used. For both EST and non-EST hits, identifier numbers were also obtained from the UniGene database.

20 The above strategy was used independently for libraries derived from macrophages and from cardiomyoblasts. By screening a relatively large number of clones (several thousand per library), single genes were identified from multiple clones from any individual library. Multiple clones covered either the same or different regions of the genes.

In the above manner, certain hypoxia-inducible genes were identified from clones only derived from the
25 cardiomyoblast library. These genes are listed in Table 1. Certain hypoxia-inducible genes were identified from clones only derived from the macrophage libraries. These genes are listed in Table 2. Certain hypoxia-inducible genes were identified from clones derived from both macrophage and myoblast libraries. These genes are listed in Table 3.

It can be seen that Table 3 contains many less genes than either Tables 1 and 2; demonstrating that these
30 cell types have large differences in the genes induced by hypoxia. Importantly, the subtracted libraries for macrophage and cardiomyoblast were constructed in parallel. Therefore, major differences in the spectrum of genes isolated from these libraries are likely to be due to differences in the starting material, rather than due to technical differences in the production of the libraries. Importantly, the genes contained

in these tables were confirmed to be hypoxia-regulated in the relevant cell type(s) by the described two-stage array hybridisation screening process.

From Table 3 it is clear that although this subset of genes was found in subtracted libraries from both hypoxic macrophages and cardiomyoblasts, the fold-induction obtained between hypoxia and normoxia, for the different tissues differs widely. For the first 5 genes in this table, the hypoxia response is greater for macrophages, whereas for the last 2 genes it is greater for cardiomyoblasts.

To test whether genes isolated only in the macrophage-derived subtracted libraries are not responsive to hypoxia in cardiomyoblast, cardiomyoblast cDNA isolated from normoxic and hypoxic cells was hybridised to an array of macrophage-derived clones. These data are presented as a scatter plot, showing normalised signal intensities in hypoxia versus normoxia, with each dot representing a single gene on the array. This plot is presented in Figure 1. A gene that is not affected by hypoxia will localise around the $y=x$ line, running diagonally through the centre of the graph. From the figure, it can be seen that most genes lie in this region, even though all the genes were responsive to hypoxia in the macrophage (Table 2). There is a subset of genes that lie beneath this region ($x>y$), representing induction of these genes by hypoxia in the cardiomyoblast.

Sequence data for the cDNA inserts of clones from our custom subtracted cDNA libraries is available. These are usually short fragments of 300-1000 bp. Some have been resequenced to obtain an accurate full insert sequence (see document "gene sequences/analysis").

Several of the genes presented in Tables 1-3 encode hypothetical proteins of unknown function and others have no database matches with protein coding sequence. The work presented here provides some functional annotation for these genes, and potential applications for the treatment of disease. Certain genes, in particular the glycolytic enzymes and transporters, have been hypothesised previously as forming part of the generic hypoxia response. The data provided herein provide solid, validating data for these hypotheses.

It was surprising to note that cells from our cultures of human ventricle-derived cells, showing a cardiomyoblast-like phenotype, do not support significant induction of the following genes: Lactate dehydrogenase A., Enolase 1, Phosphoglycerate kinase 1, Triosephosphate isomerase 1. These genes have been identified as being targets of the "ubiquitous" transcription factor HIF-1 alpha ("HIF-1: mediator of physiological and pathophysiological responses to hypoxia" *J.Appl.Physiol* 88: 1474-1480 (2000)).

Example 1b: Preparation of custom array

To confirm the findings presented in Example 1a, and to obtain more accurate and additional data, both the subtracted cDNA library clones and the IMAGE clones identified from the Research Genetics Human GeneFilters have now been fabricated by the authors into an independently produced and verified gene

array (referred to herein as the "custom gene array"), composed of PCR-amplified insert DNA. The methods used to produce this array are common in the art, but the key points are summarised below.

Clones from the subtracted cDNA library were PCR amplified as defined in Example 1a. In many cases, there were multiple cDNA clones corresponding to different regions of the same gene, and all these were
 5 represented on the custom gene array. IMAGE clones were obtained from the UK MRC HGMP Resource Centre (Hinxton, Cambridge CB10 1SB, UK) and were re-isolated as individual colonies and sequenced to verify the correct identity of the clone. In the majority of cases, the same IMAGE clone identified from the Research Genetics Human GeneFilters was selected, but in some instances these clones were not available and alternatives were selected, corresponding to the same gene.

10 Additional genes, with well-defined roles in various disease processes relevant to hypoxia, were also represented on the array, as derived from IMAGE clones. It is well established in the literature that genes with similar functions are often co-regulated at the mRNA level, as determined by microarray data clustering methods (Iyer VR *et al*, *Science*, 1999 283(5398):83-7; Eisen MB *et al* *Proc Natl Acad Sci U S A*, 1998 95(25):14863-8). This allows associations to be made between genes of unknown function (as
 15 present in the current specification) to genes of well defined function, in order to add significance to the former.

Normalisation is a key issue in array analysis. The custom gene array is a single colour type array, and contains a selection of additional IMAGE clones corresponding to genes which were empirically determined not to be affected by hypoxia and which are highly expressed in a wide range of human
 20 tissues and cell types. During data analysis, spot intensities were divided by the mean of all the reference genes shown below, each of which was present in quadruplicate on each array.

Gene	IMAGE clone Acc.
FLJ11102 fis clone PLACE1005646	AA464704
25 matrix Gla protein	AA155913
guanine nucleotide binding protein alpha stimulating 1	R43581
DKFZp434A1319	W74725
cDNA FLJ23280 fis clone HEP07194	AA669443
beta actin	(in house clone)
30 EF1a-like protein	A1817566
ribosomal protein L37a	W91881

IMAGE clone plasmid miniprep DNA was prepared and PCR amplified with flanking vector primers of the sequences GTTTTCCCAGTCACGACGTTG and TGAGCGGATAACAATTTTCACACAG. This was

then purified and concentrated by ethanol precipitation, and the presence of a single band and DNA concentration were determined by agarose gel electrophoresis and by digital imaging methods.

Purified PCR product corresponding to all the clones (IMAGE and non-IMAGE) were normalised to 0.5 mg/ ml by dilution. Arrays were fabricated onto Hybond N+ (Amersham) membranes using a
5 BioRobotics TAS arrayer (BioRobotics, Cambridge CB37LW, UK) with a 500 micron pin tool. Using 384-well source plates and a 2x2 arraying format this array was relatively low density, thereby eliminating problems of spot-to-spot signal bleed. Also the large pin size and high source plate DNA concentration improves the sensitivity of detection. Post-arraying denaturation/ neutralisation was essentially as described by Bertucci F *et al.*, 1999 (*Oncogene* 18: 3905-3912).

10 Total RNA was extracted from cells using RNeasy (Qiagen) and 7 micrograms RNA was labelled with 100 microCi 33P dCTP using 2 micrograms poly dT (10-20 mer) as primer in a reverse transcription reaction. First strand RNA was then degraded under alkaline conditions, and this was then neutralised with Tris HCl pH 8.0, and the labelled cDNA was purified using BioRad BioSpin-6 chromatography columns. Pre-hybridisation was performed in 4 ml Research Genetics MicroHyb solution supplemented with
15 10micrograms poly dA (10-20 mer) and 10 micrograms Cot-1 DNA, at 45 degrees for 2-3 hours. The cDNA was then denatured by heating and added to the pre-hybridisation, which was continued for 18-20hr. Washing steps were done as follows: 2xSSC/ 1% SDS 2x20min at 50 degrees and 0.5xSSC/ 1% SDS 10min at 55 degrees. Arrays were exposed to Amersham Low Energy phosphor screens for 24hr and scanned using a phosphorimager at 50 micron resolution. Image analysis was done using ArrayVision
20 software (Imaging Research Inc). Tab delimited data files were exported and a full analysis performed using GeneSpring software (Silicon Genetics).

Using the described methodology a dynamic range of detection of 4 logs and a sensitivity of at least 1 / 50,000 is obtained, as determined by spike doping titration experiments. Having several technical differences compared to the Research Genetics Human GeneFilters as used in the initial filing, data from
25 the custom gene array is expected to be quantitatively different.

Example 1c: Hypoxia regulation of gene expression in macrophages by exposing cells to hypoxia +/- additional signal amplification.

The transcription factor HIF-1 α , is ubiquitously present in cells and is responsible for the induction of a number of genes in response to hypoxia. This protein is considered a master regulator of oxygen
30 homeostasis (see, for example, Semenza, (1998) *Curr. Op. Genetics and Dev.* 8:588-594). Although HIF-1 α is well known to mediate responses to hypoxia, other transcription factors are also known or suspected to be involved. These include a protein called endothelial PAS domain protein 1 (EPAS1) or HIF-2 α , which shares 48% sequence identity with HIF-1 α (Tian H, *et al. Genes Dev.* 1997 11:72-82.). Evidence

suggests that EPAS1 is especially important in mediating the hypoxia-response in certain cell types, and it is clearly detectable in human macrophages, suggesting a role in this cell type (Griffiths et al., 2000, *Gene Ther.*, 7(3):255-62).

As supporting evidence for the hypoxic regulation of the genes contained within this specification,
5 adenoviral vectors were used to overexpress HIF-1 α and EPAS1 in primary human macrophages prior to exposure to hypoxia, in order to amplify the response. Because the role of these transcription factors as mediators of the hypoxia response is very well established, any further increases in the inducibility of specific genes resulting from this approach represents credible supporting evidence that those genes are responsive to hypoxia.

10 A commercially available system was used herein to produce adenoviral particles involving the adenoviral transfer vector AdApt, the adenoviral genome plasmid AdEasy and the packaging cell line Per-c6 (Crucell, Leiden, The Netherlands). The standard manufacturer's instructions were followed. Three derivatives of the AdApt transfer vector have been prepared, named AdApt ires-GFP, AdApt HIF-1 α -ires-GFP and AdApt EPAS1-ires-GFP. In these vectors, for convenience, AdApt was modified such
15 that inserted genes (i.e. HIF-1 α or EPAS1) expressed from the powerful cytomegalovirus (CMV) promoter were linked to the green fluorescent protein (gfp) marker, by virtue of an internal ribosome entry site (ires). Therefore presence of green fluorescence provides a convenient indicator of viral expression of HIF-1 α or EPAS1 in transduced mammalian cells. The control vector AdApt ires-GFP was used to allow discrimination between effects of the inserted genes (i.e. HIF-1 α or EPAS1) to that of
20 potential non-specific effects of adenoviral transduction or GFP expression. Standard subcloning methods were used to construct the adenoviral constructs as described in detail elsewhere (see co-pending, co-owned International patent application PCT/GB01/00758; Example 2).

The adenoviral transfer vectors AdApt HIF-1 α -ires-GFP and AdApt EPAS1-ires-GFP, were verified prior to production of adenoviral particles, for their ability to drive expression of functionally active HIF-1 α or
25 EPAS1 protein from the CMV promoter in mammalian cells. This was achieved by transient transfection luciferase-reporter assays as described (Boast K et al *Hum Gene Ther.* 1999 Sep 1;10:2197-208).

Using the aforementioned Introgene adenoviral system, caesium-banded, pure adenoviral particles were produced for each of the vectors AdApt ires-GFP, AdApt HIF-1 α -ires-GFP and AdApt EPAS1-ires-GFP. Following the Introgene manual, adenoviral preparations were quantitated by spectrophotometry, yielding
30 values of viral particles (VP) per milliliter.

Primary human macrophages isolated as described above, were washed and resuspended in DMEM (Gibco, Paisley, UK) supplemented with 4% fetal bovine serum (Sigma). 5×10^6 cells were plated into nine individual 10cm Primaria (Falcon) tissue culture dishes containing medium plus adenovirus as

shown below (experimental design), to give a total volume of 10 ml per plate. Two doses of adenovirus were used; 5.3×10^8 viral particles / ml (low) and 1.6×10^9 viral particles / ml (high). These amounts were chosen following a series of titration experiments. Following culture for 16 hr, during which the macrophages adhere to the plate and are infected by the adenoviral particles, the medium was removed
 5 and replaced by IMDM medium (Gibco, Paisley, UK) supplemented with 2% human AB serum. A further 24 hr period of culture was allowed prior to experimentation, to allow gene expression from the transduced adenovirus. Gene transduction was verified by visually assessing gfp expression and expression of the viral HIF-1a and EPAS1 genes was determined by real time quantitative RT-PCR using an ABI Prism 7700 TaqMan and CyberGreen protocol. For the high doses of virus, the total levels of
 10 HIF-1a or EPAS1 mRNA present in the transduced cells were increased by 10-30 fold.

For experimentation with conditions of hypoxia, identical culture dishes were divided into two separate incubators: One at 37 degrees, 5% CO₂, 95% air (=Normoxia; equivalent to 20% Oxygen) and the other at 37 degrees, 5% CO₂, 94.9% Nitrogen, 0.1% Oxygen (=Hypoxia). After 6 hours culture under these conditions, the dishes were removed from the incubator, placed on a chilled platform, washed in cold PBS
 15 and total RNA was extracted using RNeasy (Qiagen) following the manufacturer's instructions.

Experimental design

Condition	Adenovirus (type)	Adenovirus amount (low= 5.3×10^8 vp/ml high= 1.6×10^9 vp/ml)	Oxygen (%)
20			
1	none	none	20
2	AdApt ires-GFP	low	20
3	AdApt ires-GFP	high	20
4	AdApt ires-GFP	low	0.1
25	5 AdApt ires-GFP	high	0.1
6	AdApt HIF-1a-ires-GFP	low	0.1
7	AdApt HIF-1a-ires-GFP	high	0.1
8	AdApt EPAS1-ires-GFP	low	0.1
9	AdApt EPAS1-ires-GFP	high	0.1

30

RNA samples from the experimental conditions shown above were each hybridised to individual copies of the Custom gene array and processed as described earlier. To ensure reproducible data, this was repeated so each RNA sample was hybridised to 4 separate arrays. Therefore a total of 36 arrays were used for this experiment. Data analysis was done taking the mean signal of each spot from the four array
 35 replicates of each RNA sample. When displayed graphically, standard error of the mean is displayed as

the error bar. Expression values were calculated so that they represent the fold-change ratio as compared to condition#1, i.e. untreated cells.

For genes shown in Table 4 it can be seen that in cells transduced by the control adenovirus AdApt ires-GFP there is a response to hypoxia (conditions 4,5) as compared to in normoxia (conditions 2,3).
5 However this response is significantly greater when the natural hypoxia response is amplified by overexpression of HIF-1alpha from the adenovirus AdApt HIF-1a-ires-GFP (conditions 6,7). Furthermore, this effect is usually dependent on the amount of HIF1alpha overexpression (i.e. greater in condition 7 compared to 6). This same data is displayed graphically in Figure 2. It can be seen that these genes encode metallothionein proteins. One of these (Nucleotide Seq ID No. 84; Protein Seq ID No. 83)
10 is a novel member of the metallothionein family. Several metallothionein genes are known in the art to be activated by hypoxia, supporting the usefulness of this data.

For genes shown in Table 5 and Figure 3 it can be seen that in cells transduced by the control adenovirus AdApt ires-GFP there is a response to hypoxia (conditions 4,5) as compared to in normoxia (conditions 2,3). However this response is significantly greater when the natural hypoxia response is amplified by
15 overexpression of EPAS1 from the adenovirus AdApt EPAS1-ires-GFP (conditions 8,9).

In the case of the protein encoded by Seq ID No. 24, results are available independently for two separate cDNA clones representing non-overlapping regions of the same full length gene.

In the case of the protein encoded by Seq ID No. 86 (EGL nine (C.elegans) homolog 3), additional evidence is described above in support of the function of this protein. Furthermore, real time quantitative
20 RT-PCR analysis of this gene using an ABI Prism 7700 TaqMan and CyberGreen protocol, has been performed, to verify and more accurately quantitate the upregulation of EGL nine (C.elegans) homolog 3 in response to hypoxia and EPAS1 adenoviral overexpression. The main difference between the array-based and real time quantitative RT-PCR methodologies is that the latter is far more sensitive and therefore can detect expression in the off-state (here normoxia) for weakly expressed genes. This data has
25 shown an induction ratio of 819-fold for EGL nine (C.elegans) homolog 3 in response to hypoxia with additional EPAS1 expression, from RNA generated from an independent experiment. This data was normalised to beta actin.

Similarly another weakly-expressed EPAS1-induced gene, Semaphorin 4b (Seq ID No. 91/92; see additional discussion above) has been determined using real time quantitative RT-PCR methodology,
30 showing an actin-normalised induction ratio of 30.1 is found (data not shown).

For the gene shown in Table 6 and Figure 4 it can be seen that in cells transduced by the control adenovirus AdApt ires-GFP, there is a negative response to hypoxia (conditions 4,5) as compared to in normoxia (conditions 2,3). However, this response is significantly greater when the natural hypoxia response is amplified by overexpression of HIF1 alpha or EPAS1 (conditions 6,7,8,9).

5 Example 2: Differences in the hypoxia responses of resting and activated macrophages.

Macrophages accumulate at hypoxic areas in various disease states, including cancer, rheumatoid arthritis, atherosclerosis and wound healing. At these sites macrophages activation is liable to occur, such as in response to T-cell derived gamma interferon. For instance, in atherosclerotic plaques there is an accumulation of both T-cells and macrophages, and these are known to interact with one another
10 (reviewed in Lusis AJ, Atherosclerosis. Nature. 2000 Sep 14;407(6801):233-41).

It is well established that the macrophage has a significant role in the pathology of the above diseases involving hypoxia, and that most functions of the macrophage (including inflammatory functions) are greatly increased following activation. Therefore any therapeutic strategy aimed at the hypoxic macrophage, needs to also consider the effects of macrophage activation and possible cross talk between
15 the responses to macrophage activation and hypoxia.

2.1: Research Genetics Human GeneFilters

This work was carried out using Research Genetics Human GeneFilters, which contain DNA derived from clones of the IMAGE cDNA collection, representing genes of varying degrees of characterisation. A series of 6 arrays of human genes were used (GeneFilters GF200-205), potentially covering a total of
20 31,104 genes. Generally, single genes are represented only once in these arrays. However, sometimes IMAGE clones initially thought to represent separate genes, upon re-analysis were found to be different regions of the same gene. Here we have presented data for all clones individually, though they possess the same UniGene ID and gene name. An example is Hypothetical protein FLJ20037.

The methodology for Research Genetics arrays is similar in principle to that described for the array
25 screening of clones from subtracted libraries. There are several attributes to this method: Relatively small amounts of RNA can be labelled to make cDNA probes, in a single step reaction, and probes are labelled with the same chemical group (33P), so there are no errors introduced as a result of using different dyes, which may differ in stability etc. Using a Phosphorimager allows detection over a wide range of intensities (over 4 logs). Overall it is interesting to note a recent study, which has favourably re-evaluated
30 the performance of the nylon based array, as compared with the glass-based microarray method (Bertucci F *et al*, *Hum Mol Genet* 8:1715-1722 (1999)).

Experiments were done essentially as described in the Research Genetics GeneFilters protocol. Duplicate copies of each array from the same production batch, were used and hybridised in parallel with labelled RNA isolated from normoxic and hypoxic primary human macrophages. Hybridised arrays were scanned twice using a Molecular Dynamics Storm phosphorimager, and both images were analysed to ensure
 5 reproducibility. Furthermore, the experiments were repeated using the same RNA samples, but with different array lot numbers, again to ensure reproducibility.

Analysis was performed using Research Genetics Pathways software, with normalisation using the 'all data points' option. Analyses were output as spreadsheets and filtered to remove data points where the signal intensity was less than 4-fold above the general background for the experimental condition with the
 10 higher signal (hypoxia or normoxia depending on whether hypoxia causes induction or repression). Sometimes expression in the lower state was not significantly above background, and the ratio will therefore be underestimated. Ratios were calculated by normalised signal intensity in hypoxia divided by normoxia. Changes were verified visually from the original array images.

In this manner, comparisons were made between normoxia and hypoxia in resting macrophages. The
 15 whole procedure was then repeated for activated macrophages, to investigate possible differences in the response to hypoxia. It is possible that potential differences for certain genes could be correlated with changes in expression resulting from activation, prior to challenge with hypoxia. To explore this possibility, comparisons were made between resting and activated macrophages, both in normoxia. Since some of the genes we have identified as being activated by hypoxia have very low hybridisation signals in
 20 normoxia (for both resting and activated macrophages), this comparison was not possible.

We have found various patterns of gene expression changes occurring in response to hypoxia, related to the activation state of macrophages, which are presented below. Such a range of responses, specific to various subsets of genes, was not expected, and contradicts a view that the hypoxia response is a largely a generic mechanism.

25 Table 7 shows genes that are induced by hypoxia to a similar degree in resting and activated macrophages.

Table 8 shows genes that are induced by hypoxia to a greater degree in resting macrophages, as compared to activated macrophages. These data are presented illustratively in Figure 5.

Data from Table 8/Figure 5 reveals several unexpected observations.

30 A) From the final column it can be seen that macrophage activation in the absence of hypoxia, causes induction of many of these genes. This suggests that the signalling pathways resulting from activation and hypoxia might converge to a single transcriptional regulator, causing macrophage activation to pre-empt the response to subsequent hypoxia. This is exemplified most strikingly for

Interleukin 8, which is dramatically induced in response to macrophage activation, but shows no additional response to hypoxia.

- B) Genes in rows 11, 13 and 14 have no response to hypoxia following macrophage activation, though there is not a preceding large increase in expression in response to macrophage activation alone. This suggests that in the activated macrophage, the necessary signalling pathway or transcriptional regulator is not functional.
- C) Although Table 8 was produced electronically, without selecting genes based on their names, it can be seen that genes encoding proteins of the metallothionein family feature strongly.

Table 9 shows genes which are induced by hypoxia to a greater degree in activated macrophages, compared to resting macrophages. These data are presented illustratively in Figure 6.

In Table 7, there are several genes for which hypoxia/ normoxia ratios were only obtained for activated macrophages, such as Cox-2 (see row 47). For these genes, macrophage activation usually increases expression of the gene to detectable levels, thus allowing the study of subsequent changes in response to hypoxia. It is likely that these genes are not significantly expressed in resting macrophages irrespective of hypoxia, and therefore the hypoxia response is probably specific to activated macrophages.

Certain genes respond to hypoxia by decreasing mRNA expression (repression), and these genes therefore have hypoxia/normoxia ratios of < 1.0 . This phenomenon is known in the field of hypoxia, although the mechanism is obscure. Data is presented in tables 7-9, which unexpectedly shows that this hypoxia-induced repression for specific genes is not a generic process, but is dependent on the cellular context. In Table 10/ Figure 7, genes are presented that are hypoxia-repressed to a greater degree in activated (column 7) compared with resting (column 8) macrophages. Prior to any hypoxic challenge, these gene are induced to varying degrees, in response to macrophage activation (column 9), suggesting a shared mechanism for these separate responses. From Table 10, genes in rows 1-6 show that macrophage activation is necessary to obtain any response to hypoxia. In resting macrophages, these genes are not responsive to hypoxia at all.

Strikingly, Table 10/ Figure 7 shows that seven separate genes encoding chemokine proteins (Monocyte chemotactic protein 1, Macrophage inflammatory protein 1b, Monocyte chemotactic protein 3 and Small inducible cytokine A3, Monocyte chemotactic protein 2, Macrophage inflammatory protein 2a and Macrophage inflammatory protein 2 precursor) are more strongly repressed in activated macrophages as compared to resting macrophages. These genes are also among the most inducible in response to activation alone, in normoxia (column 9). These findings are of potential utility in view of the great significance of chemokines to inflammatory disease. For example, macrophage chemotactic factor 1

(Table 10, row 19) is key to the pathological role of the macrophage in atherosclerosis ("Chemokines and atherosclerosis" Reape TJ and Groot PHE, Atherosclerosis 147: 213-225, 1999).

Genes in rows 20-30 of Table 10, were not detectably expressed in resting macrophages, irrespective of hypoxia. Table 11 shows other genes that were down-regulated in response to hypoxia in macrophages.

5 Example 3: Tissue-specific hypoxia regulation of gene expression by an analysis of a series of primary human cell cultures.

Equivalent cultures of non-immortalised, non-transformed primary human cells of 10 distinct types, were cultured in either normoxia or were exposed to hypoxia for 6 hr and 18 hr, and gene expression changes were determined. To the inventors' knowledge, this is the first time that such a study has been reported.

- 10** Moreover, unlike the vast majority of information in the public domain relating to genes responsive to hypoxia, all of these cells were human and were cultured without any modifications following isolation from the human donors. By using primary cells rather than cell lines or immortalised cultures, the findings of this work more accurately represents the situation in the human body.

- Most cell types were obtained from Clonetics (distributed by BioWhittaker, Walkersville, MD) and
- 15** cultured according to the manufacturer's recommendations, unless where otherwise shown. #1: adipocyte (Clonetics CC-2568; derived from subcutaneous adult adipose tissue), #2: cardiomyocyte (Clonetics CC-2582; derived from fetal tissue; prior to experimentation cultured in minimal medium: DMEM, 4% Horse serum), #3: endothelial (TCS CellWorks ZHC-2101 human umbilical vein endothelial cells), #4: fibroblast (Clonetics CC-2511 dermal fibroblasts derived from adult tissue), #5: hepatocyte (Clonetics CC-2591,
- 20** derived from adult tissue), #6: macrophage (derived from human blood as described elsewhere in the specification), #7: mammary epithelial (Clonetics CC-2551; derived from adult tissue), #8: monocyte (derived from human blood as described elsewhere in the specification but without the 7 day differentiation culture period), #9: neuroblastoma (neuroblastoma-derived cell line SH-SY5Y), #10: renal epithelial (Clonetics CC-2556; derived from fetal tissue), #11: skeletal muscle myocyte (Clonetics CC-
- 25** 2561; derived from adult tissue). A non-primary cell type (#9) was used to represent neurons, since primary human neurons are difficult to source. Therefore a total of 11 cell types are compared. It should be noted that RNA from hepatocytes at the 16hr timepoint of hypoxia was not available for this work.

- Genes which were induced or repressed preferentially in particular cell type(s) were identified by hybridisation of the RNA samples to the custom gene array, as described in Examples 1b and 1c. Each
- 30** RNA sample was hybridised to duplicate or triplicate arrays, to ensure reproducible data, and was analysed using GeneSpring software. Data from replicate arrays were merged during analysis to generate mean values. Data normalisation was achieved per-array using the aforementioned list of control genes, such that differences in RNA labelling or hybridisation due to experimental variation were corrected by

referencing each gene to the mean value of the reference genes on the same array. Also, for each gene, expression values were obtained which represent the value in each experimental condition (e.g. macrophages 6hr hypoxia) as compared to the median of value of that gene throughout the full range of experimental conditions (i.e. from all cell types). This transformation does not alter the relative values of
 5 any gene between the different experimental conditions, and is done since there is no obvious single reference experimental condition to create ratio values. This is common in microarray data analysis.

Table 12 shows the full dataset of this analysis. From this it can be seen that certain genes respond to hypoxia differently, depending on the particular cell type. This information is valuable in identifying biological targets for the development of therapeutic and diagnostic products. Not only does it indicate a
 10 particularly significant role for these genes in the specific cell type implicated in a disease, but it also identifies that any therapeutic product is less likely to produce problematic toxicological effects. Data shown in Table 12 and the derived figures, are reproducible, and are an accurate determination of mRNA expression levels. This may be confirmed by independent means, such as quantitative real time RT-PCR.

Certain genes from Table 12 will be presented for illustration.

15 Genes with a greater response in monocytes or macrophages

Since monocytes and macrophages are similar cell types, the latter derived from the former, they will be analysed together.

Expression profiles of 11 genes showing hypoxia-induced changes in gene expression which are most pronounced in monocytes or macrophages are shown in Figures 8-18. These genes correspond to:

- 20 Seq ID:339/340 CYP1 (cytochrome P450, subfamily XXVIIIB)
- Seq ID:357/358 interleukin 1 receptor antagonist
- Seq ID:375/376 Regulator of G-protein signalling 1
- Seq ID:389/390 GM2 ganglioside activator protein
- Seq ID:405/406 Alpha-2-macroglobulin
- 25 Seq ID:475/476 Ecotropic viral integration site 2A=
- Seq ID:433/434 high affinity immunoglobulin epsilon receptor beta (CFFM4)
- Seq ID:431/432 Pleckstrin
- Seq ID:469/470 cytokine effector of inflammatory response SCYA3L
- Seq ID:79/80 Novel PI-3-kinase adapter
- 30 Seq ID:21/22 Hypothetical protein PRO0823

It will be appreciated that the majority of these genes have a known biological function in immunity/ inflammation, consistent with the known function of the monocyte/ macrophage. Further to this knowledge, this data identifies that in hypoxic disease sites where monocyte/ macrophages make up a

significant proportion of the cell types, such as in rheumatoid arthritis synovial membranes, that these genes are possible therapeutic targets.

Ecotropic viral integration site 2A (Seq ID:475/476)

For example, the gene illustrated in Figure 8, Ecotropic viral integration site 2A (Seq ID:475/476) is induced in hypoxic monocytes to a level over 25 times higher than the median expression level of this gene throughout the other cell types. This gene, of unknown function, is located on Chromosome 17q11.2 close to genes with immune functions. Presented elsewhere in this specification is data showing that expression of Ecotropic viral integration site 2A is downregulated in response to the inflammatory cytokine interferon gamma. These novel data provide evidence that Ecotropic viral integration site 2A is a novel target for inflammatory conditions involving hypoxia and monocytes.

Novel PI-3-kinase adapter Seq ID:79/80 Clone p1E9 (EST accession R62339).

Another example, in Figure 9a, is Seq ID:79/80 (EST accession R62339). It is seen that in hypoxic macrophages, this gene is expressed at 6-fold higher levels than the median expression level of this gene throughout the other cell types. Therefore, the levels of the encoded protein in hypoxic monocytes/macrophages, as found at various disease sites, are likely to be higher than in other cell types not involved in the disease process or present at the site of disease. This illuminates a novel utility of this gene as a target for the development of therapeutic products for diseases involving monocytes/ macrophages and hypoxia.

The data that led to the generation of this Figure are as follows:

20	<u>Cell type</u>	<u>Oxygen</u>	<u>Normalised expression</u> (clone p1E9 / SeqID:79/80)
	adipocyte	normoxia	1.54
	adipocyte	hypoxia 6hr	0.89
	adipocyte	hypoxia 18hr	1.48
25	cardiomyocyte	normoxia	1.18
	cardiomyocyte	hypoxia 6hr	1.80
	cardiomyocyte	hypoxia 18hr	1.53
	endothelial	normoxia	0.68
	endothelial	hypoxia 6hr	0.82
30	endothelial	hypoxia 18hr	0.60
	fibroblast	normoxia	0.60
	fibroblast	hypoxia 6hr	0.64
	fibroblast	hypoxia 18hr	0.73
	hepatocyte	normoxia	0.92
35	hepatocyte	hypoxia 6hr	1.62
	macrophage	normoxia	4.20
	macrophage	hypoxia 6hr	3.97
	macrophage	hypoxia 18hr	6.19
	mammary epithelial	normoxia	0.25
40	mammary epithelial	hypoxia 6hr	0.42

	mammary epithelial	hypoxia 18hr	0.18
	monocyte	normoxia	2.33
	monocyte	hypoxia 6hr	3.63
	monocyte	hypoxia 18hr	5.01
5	neuroblastoma	normoxia	0.93
	neuroblastoma	hypoxia 6hr	0.80
	neuroblastoma	hypoxia 18hr	0.85
	renal epithelial	normoxia	0.57
	renal epithelial	hypoxia 6hr	0.61
10	renal epithelial	hypoxia 18hr	0.61
	skeletal myocyte	normoxia	1.58
	skeletal myocyte	hypoxia 6hr	1.37
	skeletal myocyte	hypoxia 18hr	1.17

- 15 To substantiate the array-based data, the same RNA samples were examined by real time quantitative RT-PCR. The advantages of this method are that it is more sensitive and because two gene-specific primers are used, the data will be more specific to the gene in question.

RNA from the above samples (except for the hepatocyte RNA which was unavailable) was Dnase I-treated prior to reverse transcription to remove possible contaminating genomic DNA and was reverse transcribed using an oligo dT₍₁₅₎ primer and Superscript II reverse transcriptase. These samples were used as template for PCR reactions using primers specific to EST accession R62339 or to beta-actin. Primer sequences were as follows:

Novel PI-3-kinase adapter Seq ID:79/80 Clone p1E9 (EST accession R62339).

Forward Primer 5' GCC CTT AGT TTT TCA CTT CTT CGT 3'

- 25 Reverse Primer 5' CCT TAA GAT CCA TTC TCA TTG CTG AT 3'

Beta Actin

Forward Primer 5' GCC CTG AGG CAC TCT TCC A 3

Reverse Primer 5' GCG GAT GTC CAC GTC ACA 3'

- All RT-PCR reactions were performed using an ABI Prism 7700 Sequence Detector system. For each Q-PCR run, a master mix was prepared with 2x SYBR Green I master mix (Applied Biosystems) and primers at 5 μ M. Two microlitres of respective diluted cDNA were added to PCR master mixture, amounting to 25 μ L. The thermal cycling conditions comprised 50°C for 2 minutes, 95°C for 10 minutes, 40 cycles at 95°C for 15 seconds, and 60°C for 1 minute. PCR reactions were set up in 96 well format with duplicate amplifications for each data point including 8 serial cDNA dilutions (0.2, 0.1, 0.05, 0.025, 0.01, 0.005, 0.001 and 0.0001) of macrophage treated with 18 hours hypoxia to compose a standard curve, a no template control, no amplification control lacking reverse transcriptase, and each cDNA sample at a dilution value of 0.1. The experiment for the novel PI3K adapter was carried out in triplicate for reproducibility which were later determined by linear regression analysis. Data was analysed with

necessary adjustment of the default baseline and threshold line using ABI Prism 7700 software. The C_t value, an important raw data for each sample, was calculated as the cycle number at which the ΔR_n crosses the baseline. For each run, a standard curve was constructed by plotting a graph with mean C_t values from 8 data points from standard sample against log input of the corresponding dilution values with a best fit trend line. From the trend line, the formula ' $y=mx+c$ ' was created according to the y-intercept and slope of standard curve which then were used for calculating the log input amount of the experimental cDNA samples, as related to the calibration sample. Data for the Novel PI-3-kinase adapter was normalized to that of beta-actin to correct for potential differences in efficiency of cDNA synthesis between the RNA samples.

- 10 From the TaqMan data the specificity to monocytes and macrophage found from the array data is confirmed and found to be even more pronounced (see Figure 9b). The data presented in the Figure are listed below. In the data listed below, the normalized expression values are multiplied by 1000 for clarity.

Cell type	Oxygen	Normalised expression (clone p1E9 / SeqID:79/80)
15 adipocyte	normoxia	0.050
adipocyte	hypoxia 6hr	0.007
adipocyte	hypoxia 18hr	0.015
cardiomyocyte	normoxia	0.163
cardiomyocyte	hypoxia 6hr	0.037
20 cardiomyocyte	hypoxia 18hr	0.222
endothelial	normoxia	3.093
endothelial	hypoxia 6hr	0.059
fibroblast	normoxia	0.527
fibroblast	hypoxia 6hr	0.043
25 fibroblast	hypoxia 18hr	0.037
macrophage	normoxia	404.593
macrophage	hypoxia 6hr	503.026
macrophage	hypoxia 18hr	1162.056
mammary epithelial	normoxia	0.026
30 mammary epithelial	hypoxia 6hr	0.068
mammary epithelial	hypoxia 18hr	0.112
monocyte	normoxia	565.471
monocyte	hypoxia 6hr	657.465
monocyte	hypoxia 18hr	979.048
35 neuroblastoma	normoxia	8.482
neuroblastoma	hypoxia 6hr	7.104
neuroblastoma	hypoxia 18hr	4.707
renal epithelial	normoxia	17.898
renal epithelial	hypoxia 6hr	9.831
40 renal epithelial	hypoxia 18hr	10.929
skeletal myocyte	normoxia	0.930
skeletal myocyte	hypoxia 6hr	0.638
skeletal myocyte	hypoxia 18hr	1.627

There are several technical reasons why the results from the array-based data might be more pronounced in the Taqman results - the lower sensitivity of the array-based method means that genes which are not expressed will be detected as a background signal. Also the array method is more likely to suffer from cross-hybridisation between similar genes.

- 5 The TaqMan data illustrates dramatically the concept that the hypoxia response is not just a generic response found in all cell types, relating to generic cell processes such as metabolism.

Database searches for gene sequences showing identity with IMAGE clone acc:R62339 reveal that there are no matching human sequences of any type other than ESTs. This includes full length cDNAs, truncated cDNAs, gene sequences from chromosomal data or hypothetical protein gene sequences.

- 10 Therefore the human gene represented by IMAGE clone acc:R62339 is a novel human gene.

- Although this human EST is unannotated, by comparison with mouse sequence data (acc AF293806), it appears likely to encode a novel human Phosphoinositol 3-kinase (PI3-kinase) adapter molecule, homologous to the recently described mouse gene, BCAP. This class of molecule, involved in intracellular signalling, have been shown to have utility as a drug target (see Stein RC *et al*, "PI3-kinase inhibition: a target for drug development" *Mol Med Today*. 2000 Sep;6(9):347-57). PI3-kinases are key to many cellular processes relevant to human disease, including proliferation, apoptosis and inflammation. The data presented for the gene encoded by Seq ID:79/80 provides evidence that the encoded protein is a novel drug target in humans, specifically targeting monocyte/ macrophages at hypoxic disease sites.

- In the publication relating to murine BCAP, the protein is identified as an adapter molecule connecting the non-receptor protein tyrosine kinase Syk to the p85 subunit of PI3-kinase, and therefore to the pivotal signalling pathways centred around PI3-kinase (Okada T *et al* "BCAP: the tyrosine kinase substrate that connects B cell receptor to phosphoinositide 3-kinase activation." *Immunity*. 2000 13:817-27). Although, in this report, Syk is acting as the intracellular signalling component of the B cell antigen receptor, which is present exclusively on B-cells, Syk has been shown to initiate intracellular signalling from other cell surface receptors which are expressed on macrophages, including the Fc gamma receptor, the chemokine receptor CCR5 and macrophage-expressed CD8 (Darby C *et al* "Stimulation of macrophage Fc gamma RIIIA activates the receptor-associated protein tyrosine kinase Syk and induces phosphorylation of multiple proteins including p95Vav and p62/GAP-associated protein". *J Immunol*. 1994 152:5429-37) (Kedzierska K *et al* "FcgammaR-mediated phagocytosis by human macrophages involves Hck, Syk, and Pyk2 and is augmented by GM-CSF." *J Leukoc Biol*. 2001 Aug;70(2):322-8.), (Ganju RK *et al* "Beta-chemokine receptor CCR5 signals through SHP1, SHP2, and Syk." *J Biol Chem*. 2000 275:17263-8.), (Lin TJ *et al* "Activation of macrophage CD8: pharmacological studies of TNF and IL-1 beta production." *J Immunol*. 2000 164:1783-92.).

Indeed, syk has been validated as target in macrophages to inhibit inflammatory activities of this cell type (Stenton GR et al "Aerosolized Syk antisense suppresses Syk expression, mediator release from macrophages, and pulmonary inflammation." *J Immunol.* 2000 Apr 1;164(7):3790-7.).

- Additional to the finding that the probable human orthologue of the adapter molecule BCAP is
 5 preferentially hypoxia-induced in human monocytes/ macrophages, we also find from data generated by the custom array, that the protein acting immediately upstream of BCAP (i.e. Syk) is also regulated by hypoxia in this novel cell type specific manner, greatly increasing the biological significance of the original finding (see Figure 9c). The data used to generate this Figure are presented below for clarity.

<u>Cell type</u>		<u>Oxygen</u>	<u>Normalised expression</u>
			<u>(of syk)</u>
10	adipocyte	normoxia	2.6573591
	adipocyte	hypoxia 6hr	1.499927
	adipocyte	hypoxia 18hr	1.1115488
15	cardiomyocyte	normoxia	0.8357341
	cardiomyocyte	hypoxia 6hr	2.161058
	cardiomyocyte	hypoxia 18hr	0.90880114
20	endothelial	normoxia	0.60265505
	endothelial	hypoxia 6hr	0.56874704
	endothelial	hypoxia 18hr	0.43321633
	fibroblast	normoxia	0.8542026
	fibroblast	hypoxia 6hr	0.7657573
	fibroblast	hypoxia 18hr	0.784982
25	hepatocyte	normoxia	0.5238476
	hepatocyte	hypoxia 6hr	0.8465495
	macrophage	normoxia	4.272981
	macrophage	hypoxia 6hr	6.144931
	macrophage	hypoxia 18hr	10.278416
30	mammary epithelial	normoxia	1.1023632
	mammary epithelial	hypoxia 6hr	2.7382789
	mammary epithelial	hypoxia 18hr	0.7985004
	monocyte	normoxia	6.052118
	monocyte	hypoxia 6hr	8.6809225
	monocyte	hypoxia 18hr	11.58468
35	neuroblastoma	normoxia	1.0230793
	neuroblastoma	hypoxia 6hr	1.089154
	neuroblastoma	hypoxia 18hr	0.7689335
	renal epithelial	normoxia	0.88565326
	renal epithelial	hypoxia 6hr	1.2609364
	renal epithelial	hypoxia 18hr	0.6242461
40	skeletal myocyte	normoxia	1.3959162
	skeletal myocyte	hypoxia 6hr	0.91255134
	skeletal myocyte	hypoxia 18hr	0.64795935

- 45 In summary, we have shown here that a novel human gene encoding a predicted signalling protein relevant to human disease is activated by hypoxia, specifically in monocytes and macrophages. This data

is validated by non-array based means. Furthermore, we identify the protein immediately upstream of this signalling system as being co-regulated in this manner too. Therefore the human PI3-kinase adapter encoded by IMAGE clone acc: R62339 and the non-receptor tyrosine kinase Syk are both identified here for the first time as therapeutic targets for diseases involving hypoxic macrophages, including
5 Rheumatoid arthritis, chronic occlusive pulmonary disease, atherosclerosis and cancer. Because both genes are preferentially expressed in hypoxic macrophages, toxicity effects of therapeutic products directed at the encoded proteins are likely to be limited.

As discussed in detail above, fragments and functional equivalents of the PI-3-kinase adapter protein represented in Seq ID:79/80 and other equivalent proteins are included within the present invention, in
10 addition to ligands that bind specifically to these proteins. Furthermore, the invention also embraces purified and isolated nucleic acid molecules encoding these proteins, fragments and functional equivalents, vectors containing such nucleic acid molecules and host cells transformed with these vectors.

Regulator of G-protein signalling 1 (Seq ID:375/376)

Another intracellular signalling protein, Regulator of G-protein signalling 1 (RGS1; Seq ID:375/376), in
15 shown in Figure 10. Here the expression levels in the hypoxic monocyte is 30-fold higher than the median expression level of this gene throughout the other cell types. The function of this protein is to negatively regulate G protein signalling pathways, and inhibit chemokine-induced cell migration of immune cells (Moratz C et al *J Immunol*, 2000 164:1829-38 and Denecke B et al *J Biol Chem*. 1999 274:26860-8.).

Our data suggests that this gene is preferentially expressed in macrophages, consistent with the findings
20 of Denecke B et al (*J Biol Chem*. 1999 274:26860-8.). Our novel finding that expression is even further enhanced by hypoxia illuminates a mechanism by which cell migration is inhibited in hypoxia, leading to an accumulation of these cells at pathological sites of hypoxia. This mechanism is novel and distinct to other mechanisms proposed in the art to explain this key aspect of hypoxia and inflammation (for example: Grimshaw MJ et al "Inhibition of monocyte and macrophage chemotaxis by hypoxia and
25 inflammation--a potential mechanism." *Eur J Immunol*. 2001 31:480-9).

Furthermore, Figure 10 shows that Regulator of G-protein signalling 1 is upregulated during differentiation of monocytes to macrophages, with significance to changes in cell motility. This discovery therefore provides that inhibitors of RGS1 have utility in increasing the motility of macrophages that are used for cell-based therapies. Accordingly, one embodiment of this aspect of the invention provides for
30 the use of an inhibitor of RGS1 in therapy, by increasing the motility of macrophage cells.

GM2 ganglioside activator protein

The gene shown in Figure 11, GM2 ganglioside activator protein, was originally characterised as a lysosomal co-factor required for degradation of gangliosides. It has been proposed to have alternative

roles as a secreted protein, and can bind and inhibit the actions of the inflammatory mediator, platelet activating factor (Rigat B et al *Biochem Biophys Res Commun.* 1999 258:256-9.).

Our novel finding, presented in Figure 11, shows that GM2 ganglioside activator protein is induced by hypoxia, preferentially in macrophages, suggesting an influence on the inflammatory functions of the
5 macrophage in hypoxia.

In Figures 15-18, genes are shown which are expressed preferentially in the monocyte/ macrophage, but which are *decreased* in expression in response to hypoxia. Being expressed at highest levels in the monocyte/ macrophage, these genes are more likely to be significant to the biological functions of this cell type.

10 Interleukin 1 receptor antagonist (Seq ID:357/358)

In Figure 15, the gene interleukin 1 receptor antagonist (Seq ID:357/358) is seen to be down-regulated by hypoxia in the macrophage. Since the function of the encoded protein is anti-inflammatory, then down-regulation of this gene would be expected to have a pro-inflammatory effect. Therefore, corrective expression of the gene, would be expected to produce therapeutic effects in inflammatory disorders
15 involving macrophages and hypoxia, such as Rheumatoid Arthritis (Hollander AP et al. *Arthritis Rheum.* 2001 44:1540-4). This correlates with effects seen from the application the drug Anakrina / KineretTM developed by Amgen. This supports the applicability of the genes disclosed herein as novel targets for therapeutic products.

The example of gene interleukin 1 receptor antagonist also provides good exemplification of the concept
20 that different cell types respond to hypoxia differently. Here, not only are there quantitative differences, but also qualitative differences in that this gene is *down-regulated* by hypoxia in macrophages, but *up-regulated* by hypoxia in several other cell types, such as renal epithelial cells (see Figure 15). Such findings are not documented in the art.

The dataset of Table 12 also contains genes which are induced preferentially in monocyte/ macrophages
25 and also in some but not all other cell types tested. Several of these genes are present as multiple clones on the gene array, giving separate data, therefore adding extra confidence to the conclusions. These genes, presented in Figures 19-28 correspond to:

SeqID:313/314 adipophilin

SeqID:163/164 Hypothetical protein FLJ13511

30 SeqID:267/268 Osteopontin

SeqID:17/18 Hematopoietic Zinc finger protein

SeqID:137/138 CYP1B1

SeqID:325/326 CYP1B1

It will also be seen that in the case of CYP1B1 (clones p1F16 and p1E3) the hypoxia response in monocytes / macrophages is qualitatively different to the other cell types tested, in that the gene is up-regulated rather than down-regulated in response to hypoxia.

Genes with a greater response in endothelial cells

- 5 The dataset of Table 12 also contains genes which are induced preferentially in endothelial cells, a cell type key to the process of angiogenesis, in response to hypoxia. These genes are as follows, and are presented in Figures 29-31:

SeqID:205/206 Hypothetical protein FLJ22690

SeqID:65/66 cDNA DKFZp586E1624

- 10 SeqID:197/198 EST

Genes with a greater response in hepatocytes

- The dataset of Table 12 also contains genes which are induced preferentially in hepatocytes, in response to hypoxia. These genes are presented in Figures 32a and 33-38. It is noted that most of these genes, including hqp0376, encode proteins of the metallothionein family. Furthermore, close inspection of these
15 data reveals that the fold induction in hypoxia compared to normoxia for monocyte/ macrophages are very high, though the absolute levels of expression are below that of hepatocytes.

SeqID:85/86 EGL nine (C.elegans) homolog 3

SeqID:83/84 Novel Metallothionein

SeqID:337/338 Hypothetical protein.hqp0376 (a metallothionein)

- 20 SeqID:265/266 Metallothionein 2A

SeqID:243/244 Metallothionein 1G

SeqID:141/142 Heparin antimicrobial peptide

SeqID:239/240 Metallothionein 1H

EGL nine (C.elegans) homolog 3

- 25 As described above, it has been discovered that a polypeptide encoded by a gene identified from the EST recited in SEQ ID No 86, having the Protein accession number BAB15101 (encoded by Homo sapiens cDNA: FLJ21620 fis, clone COL07838 Nucleotide accession AK025273) is regulated by hypoxia. Other public domain sequences corresponding to this gene include Homo sapiens cDNA: FLJ23265 fis, clone COL06456 Nucleotide accession AK026918. Accordingly, when referring in the present specification to
30 the EST recited in SEQ ID No 86, it is intended that these gene and protein sequences are also embraced. This gene was identified using Research Genetics Human GeneFilters arrays, which contain an EST

corresponding to the gene (accession number R00332). The gene is now termed EGL nine (*C.elegans*) homolog 3.

There are no reports that describe the function of this human gene. However, a high degree of amino acid homology is observed between the protein encoded by this gene, and a rat protein called "Growth factor responsive smooth muscle protein" or "SM20" (Nucleotide accession U06713; Protein accession A53770). An alignment of single letter amino acid sequences is shown below. Over the highlighted region there is 97% amino acid similarity and 96% amino acid identity.

10	A53770	(1)	MTLSRRGFLSFLPGLRPPRRWLRI SKRGPPTSHWASPALGGRTLHYSCR	
	BAB15101	(1)	-----	
		51		100
	A53770	(51)	SQSGTPFSSEFQATFPFAAKVARGPWLPQVVEPPARLSASPLCVRSGQA	
	BAB15101	(1)	-----	
		101		150
15	A53770	(101)	LGACTLGVPRLGSVSEMP LGHTIMRLDLEKIALEYIVPCLHEVGFCYLDNF	
	BAB15101	(1)	-----MPELGHTIMRLDLEKIALEYIVPCLHEVGFCYLDNF	
		151		200
	A53770	(151)	LGEVVGDCVLERVKQLHYNGALRDGQLAGPRAGVSKRHLRGDQITWIGGN	
	BAB15101	(35)	LGEVVGDCVLERVKQLHCTGALRDGQLAGPRAGVSKRHLRGDQITWIGGN	
20		201		250
	A53770	(201)	EEGCEAINFLSLIDRLVLYCGSRLGKYVVKERSKAMVACYPGNGTGYVR	
	BAB15101	(85)	EEGCEAISFLSLIDRLVLYCGSRLGKYVVKERSKAMVACYPGNGTGYVR	
		251		300
	A53770	(251)	HVDNPNGDGRCITCIYYLNKNWDAKLHGGVLRIFPEGKSEFADVEPIFDR	
25	BAB15101	(135)	HVDNPNGDGRCITCIYYLNKNWDAKLHGGILRIFPEGKSEFADVEPIFDR	
		301		350
	A53770	(301)	LLFSWSDRRNPHEVOPSYATRYAMTVWYFDAEERAEAKKKFERNLTRKTES	
	BAB15101	(185)	LLFSWSDRRNPHEVOPSYATRYAMTVWYFDAEERAEAKKKFERNLTRKTES	
		351		
30	A53770	(351)	AI AKD	
	BAB15101	(235)	AI TED	

The high degree of amino acid similarity suggests that the human protein BAB15101 has an equivalent biochemical function to the rat protein A53770 ("Growth factor responsive smooth muscle protein" or "SM20"). Recent publications have shown that SM20 functions to promote apoptosis in neurons (Lipscomb *et al.*, *J Neurochem* 1999; 73(1):429-32; Lipscomb *et al.*, *J Biol Chem* 2000 Nov 1; [epub ahead of print]). Significantly, SM20 has been shown to be expressed at high levels in the heart (Wax *et al.*, *J Biol Chem* 1994; 269(17): 13041-7).

It has also been discovered that a polypeptide encoded by a gene identified from the EST recited in SEQ ID No 90, having the Protein accession number CAB81622, is regulated by hypoxia. The encoding human gene has been annotated in the UniGene database as "Similar to rat smooth muscle protein SM-

20"; the nucleotide sequence is contained within the nucleotide accession AL117352. More recently, a longer fragment of this gene has been cloned, named clorf12, or EGLN1 (Nucleotide accession AAG34568; Protein accession AAG34568). Accordingly, when referring in the present specification to the EST recited in SEQ ID No 90, it is intended that these gene and protein sequences are also embraced.

5 This distinct human gene, encoding a protein related to SM20 and EGLN3 (BAB15101), is also induced in response to hypoxia. This gene was identified using Research Genetics Human GeneFilters arrays, which contain an EST corresponding to the gene (accession number H56028).

Independently to this, a fragment of this gene has been cloned from a cDNA library derived from hypoxic human cardiomyoblasts, and it has been shown that the gene is increased in expression in response to
10 hypoxia in this cell type (see Table 1 herein; penultimate row). The nucleotide sequence of this cDNA fragment is referred to herein as SEQ ID No 90a.

In the light of this novel discovery reported herein that these human equivalents of SM20 are induced by hypoxia, it is herein proposed that in cardiac ischaemia, the resulting apoptosis is due at least in part, to increased expression of these genes. The therapeutic modulation of the activity of EGLN3 (BAB15101),
15 clorf12 (AAG34568), CAB81622, SM20 and other equivalent proteins and encoding genes therefore provides a novel means for the treatment of myocardial ischaemia, through the alteration of the propensity of myocardial cells to undergo apoptosis. For example, a suitable treatment may involve altering the susceptibility of ischaemic myocardial tissue to subsequent reperfusion and re-oxygenation, or may involve modulating the susceptibility of chronic ischaemic myocardial tissue (including forms of
20 angina) to later more severe ischaemia, which would result in myocardial infarction. It is submitted that, by way of analogy, cerebral ischaemia may be treated using the same principle.

Although the Applicant does not wish to be bound by this theory, the downstream effects of SM20 and related genes such as EGLN3 (BAB15101), clorf12 (AAG34568), and CAB81622, namely, apoptosis and angiogenesis might be explained as follows. The apoptotic effect of NGF withdrawal may be
25 mediated by induction of the hypoxia pathway, but may be an aspect of the supposed involvement of the HIF protein in the stress response. HIF1 α is induced by reactive oxygen species (see Richard et al. J Biol Chem 2000 Sep 1;275(35):26765-71). This could, in turn, be mediated by over-load of the proteosomal pathway for HIF1 α degradation and the consequent accumulation of undegraded HIF1 α . Accordingly, it is considered that modulation of SM20 and the related genes EGLN3 (BAB15101), clorf12
30 (AAG34568), and CAB81622 may have applications in the treatment of diseases resulting from disturbances in proteosome function, such as prion diseases and other neuro-degenerative diseases.

These data provide the first connection between these related genes and the physiological response to hypoxia. Recently published research papers have identified that the protein products of these genes can

act as proline-hydroxylases (see Bruick RK et al Science. 2001 294:1337-40 and Epstein AC et al Cell. 107:43-54). This is consistent with our observations that certain proline hydroxylases are induced in response to hypoxia and the genes EGLN1 and EGLN3 are part of the hypoxia response. For example, two genes encoding proline hydroxylases have been identified herein as being increased in expression in response to hypoxia (proline 4-hydroxylase, alpha polypeptide I; SeqID: 231/232, proline 4-hydroxylase, alpha polypeptide II; SeqID: 349/ 350). This identified a functional significance of proline hydroxylation as a response to hypoxia.

Proline hydroxylase leads to degradation of HIF1 α in normoxia (HIF regulates its own degradation – feedback). Hydroxylated HIF1 α + VHL leads to ubiquitination and consequent degradation of HIF1 α by proteasome. The activity of the prolyl hydroxylase is O₂-dependent, so under conditions of hypoxia, HIF1 α is not hydroxylated efficiently and is stabilised. HIF1 α protein thus accumulates to a high level. The hypoxia-induction of the prolyl hydroxylase ensures that when O₂ concentration returns to normal, there is sufficient enzyme available to target this high level of HIF1 α efficiently for rapid degradation.

Degradation of HIF1 α is dependent on HIF1-induced transcription (i.e. is hypoxia inducible). Berra et al (FEBS Lett 2001 Feb 23;491(1-2):85-90) raises the specific hypothesis of an unknown hypoxia-inducible factor which targets HIF1 α for proteosomal degradation. It appears reasonable to propose that this factor will clearly be hypoxia-inducible, to ensure that a rapid and effective constraint on the hypoxic response would operate on return to normoxia. It now appears as if the genes EGLN1 and EGLN3 form part of this mechanism.

It is also hypothesised that SM20 and the related genes EGLN3 (BAB15101), c1orf12 (AAG34568), and CAB81622 may act as tetramers. Known prolyl hydroxylases such as prolyl 4-hydroxylase (P4H) are known to act as tetramers of two alpha subunits and two beta subunits. SM20 and the related genes exhibits high similarity to the alpha subunit of P4H and it therefore seems likely that SM20 and the related genes are likely to have a binding partner that is equivalent to the beta subunit of P4H.

SM20 has been shown to bind to the transcription factor HIF1 α , and shares a low level homology with a p53 binding protein. P53 is a transcription factor that is known to be involved in apoptosis. Accordingly, it is proposed that in addition to binding to HIF1 α , SM20 and the related genes EGLN3 (BAB15101), c1orf12 (AAG34568), and CAB81622 may also bind and modify other transcription factors that are involved in the hypoxic response such as EPAS and HIF3A, or other transcription factors such as p53 and thereby influencing apoptosis. This aspect of the invention thus provides dimer and tetrameric forms of the EGLN3 (BAB15101), c1orf12 (AAG34568), and CAB81622 proteins, preferably complexed with a protein selected from the group consisting of HIF1 α , p53 and a protein binding partner that is equivalent to the beta subunit of P4H. Preferably, such dimers and tetramers are heterodimers/heterotetramers.

To provide further evidence that these related genes are a significant part of the hypoxia response additional expression data is presented here.

Expression profiles for these two genes will be displayed with pre-chip normalisation to correct for differences in RNA labelling etc, but within each gene no further normalisation is done (per-gene
5 normalisation), so the relative absolute expression levels of the two genes can be compared and Y-axis units between separate graphs from the same experiment are comparable. These graphs are presented as Figures 32b (c1orf12) and 32c (EGLN3).

It can be seen from these Figures that both genes (c1orf12 and EGLN3) are inducible in response to hypoxia in macrophages whether activated by gamma interferon and lipopolysaccharide or if de-activated
10 by treatment with interleukin-10. In macrophages the absolute expression level of C1orf12 appears to be higher than EGLN3.

There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show herein that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell types.

15 From Figures 32a and 32d and the data presented below, differing expression profiles of the two related genes c1ORF12 and EGLN3 are apparent throughout the 11 tested cell types, though C1orf12 is generally expressed at higher levels than EGLN3.

Cell type	Oxygen	mRNA expression	
		(c1ORF12 SeqID:89/90)	(EGLN3 SeqID:85/86)
20 adipocyte	normoxia	0.0075	0.0033
adipocyte	hypoxia 6hr	0.0091	0.0027
adipocyte	hypoxia 18hr	0.0182	0.0025
cardiomyocyte	normoxia	0.0067	0.0019
25 cardiomyocyte	hypoxia 6hr	0.0381	0.0023
cardiomyocyte	hypoxia 18hr	0.0201	0.0026
endothelial	normoxia	0.0198	0.0019
endothelial	hypoxia 6hr	0.0583	0.0033
endothelial	hypoxia 18hr	0.0397	0.0026
30 fibroblast	normoxia	0.0119	0.0032
fibroblast	hypoxia 6hr	0.0260	0.0046
fibroblast	hypoxia 18hr	0.0235	0.0040
hepatocyte	normoxia	0.0075	0.0080

	hepatocyte	hypoxia 6hr	0.0074	0.0146
	macrophage	normoxia	0.0033	0.0008
	macrophage	hypoxia 6hr	0.0083	0.0018
	macrophage	hypoxia 18hr	0.0058	0.0021
5	mammary epithelial	normoxia	0.0065	0.0014
	mammary epithelial	hypoxia 6hr	0.0137	0.0055
	mammary epithelial	hypoxia 18hr	0.0144	0.0065
	monocyte	normoxia	0.0027	0.0006
	monocyte	hypoxia 6hr	0.0084	0.0014
10	monocyte	hypoxia 18hr	0.0080	0.0016
	neuroblastoma	normoxia	0.0344	0.0011
	neuroblastoma	hypoxia 6hr	0.1085	0.0013
	neuroblastoma	hypoxia 18hr	0.0551	0.0020
	renal epithelial	normoxia	0.0275	0.0046
15	renal epithelial	hypoxia 6hr	0.0560	0.0046
	renal epithelial	hypoxia 18hr	0.0395	0.0096
	skeletal myocyte	normoxia	0.0088	0.0029
	skeletal myocyte	hypoxia 6hr	0.0277	0.0035
	skeletal myocyte	hypoxia 18hr	0.0245	0.0038

20

For instance, in the hypoxic hepatocyte (6hr) the normalised expression values of EGLN and c1orf12 are 0.015 and 0.0074 respectively, i.e. EGLN being the dominant gene. In contrast, in the neuroblastoma cell line SH-SY5Y, the normalised expression values of EGLN and c1orf12 after 6hr hypoxia are 0.0012 and 0.108 respectively, i.e. c1orf12 being the dominant gene by a large margin. This data demonstrates that

25 c1ORF12 and EGLN3 are not constitutively expressed at an equal amount in different tissues indicating specificity of function. Therefore, it is considered that therapeutic products may be developed based on this data, with the goal of modulating proline hydroxylation of target proteins (such as HIF1alpha) in specific tissues, based on the differing expression profile of c1ORF12 and EGLN3 in those tissues.

In Example 1b herein, genes were identified from a custom array, which give a greater induction in

30 macrophages (by a factor of at least 1.5) when hypoxia is augmented by over-expression of HIF1alpha or EPAS from an adenovirus. The data from the HIF/ EPAS over-expression work is presented herein in Example 1c, but specifically relating to c1ORF12 and EGLN3 is summarised in Figures 32e and 32f. From this data it is apparent that EGLN3/ FLJ21620 is c1.COL07838 but not c1ORF12 is increased in expression by the transcription factor EPAS1 but not HIF1alpha. This is apparent by comparing

experimental condition 9 (hypoxia with EPAS overexpression; expression value=3.48) to that of 5 (hypoxia without EPAS overexpression; expression value= 1.65). This adds valuable information about the mechanism of regulation of the gene encoding EGLN3.

To confirm this data the RNA samples for experimental conditions 1,3,5,7,9 (corresponding to the high
5 dose of adenovirus) were also measured using a different array-based methodology- the AffyMetrix GeneChip. The results of this experiment are presented in Figures 32g and 32h.

Functional Characterisation of EGL nine (C.elegans) homolog 3 role in the induction of Cardiomyocyte apoptotic cell death

EGLN3 has been cloned into pONY8.1 and Smart2.IRES.GFP equine infectious anaemia virus (EIAV)
10 vectors, and AdCMV.TRACK.GFP (AdenoQuest) adenoviral genome vectors (see co-owned co-pending International patent application PCT/GB01/00758). These vectors have been used in "gain-of-function" studies in which EGLN3 has been overexpressed in order to elucidate corresponding protein function. Human embryo kidney (HEK 293T) and dog osteosarcoma (D17) cell lines have been used in transient plasmid transfection experiments to confirm EGLN3 expression from viral vector genomes. Rat
15 cardiomyocyte cell line (H9C2) and primary human neonatal cardiomyocytes (PHNC) (BioWhittaker, CC2582) have been used in viral transduction experiments to determine the biological activity of EGLN3. In all cell types, expression of EGLN3 has been followed by combinations of immunofluorescence, Western blotting and TaqMan quantitative PCR. Immunofluorescence and Western blotting employ an antibody specific for the FLAG epitope engineered into the 3' terminus of EGL nine (C.elegans) homolog
20 3 (Sigma, F3165). TaqMan quantitative PCR utilises the SYBR Green method (Applied Biosystems).

Western blotting has confirmed the transient expression of EGLN3 from an EIAV genome construct in HEK 293T (expected size approx 717 bp, 26 Kda). Immunofluorescence has localised transient expression of EGL nine (C.elegans) homolog 3 from EIAV expression construct in HEK293T to the cytoplasm. Expression of EGL nine (C.elegans) homolog 3 is elevated after 4 hours exposure to hypoxic
25 conditions (0.1% (v/v) oxygen), when compared to expression observed under normoxia (20% (v/v) oxygen) (see Figure 32i). TaqMan primers have been designed and optimised for the initial measurement of EGL nine (C.elegans) homolog 3 expression in EIAV or Adenovirus transduced H9C2 and PHNC (Forward: TCATCGACAGGCTGGTCCTC; Reverse: GTTCCATTTCCTGGATAGAA). All findings at the RNA level are corroborated by immunofluorescence and Western blotting analyses at the protein
30 level.

EIAV transduction of H9C2 and PHNC has been optimised with constructs containing green fluorescence protein (GFP) and LacZ reporter genes, using the VSVg envelope and a range of MOI between 10 and 100. GFP results were scored by fluorescence microscopy, while LacZ transductants were identified

through the assay of β -galactosidase activity. An MOI of 50 transduced approximately 50% of the cell population.

EGLN3 is predicted to have pro-apoptotic activity in cardiomyocytes. Early, Mid and late phase apoptosis are characterised by translocation of membrane phospholipid phosphatidylserine (PS) from the inner face
5 of the plasma membrane to the cell surface, activation of specific proteases (caspases) and fragmentation of DNA, respectively (Martin, S.J., et al., J. Exp. Med. 1995, 182, 1545-1556; Alnemri, E.S., et al., J. Cell. Biochem. 1997, 64, 33-42; Wylie, A.H., et al., Int. Rev. Cytol. 1980, 68, 251-306). Translocation of PS has been identified through use of ApoAlert kit (Clontech; K2025-1), which employs FITC-labelled antibodies to detect surface expression of the PS, Annexin V. Caspase activity has been followed using
10 the homogeneous fluorimetric caspase assay (Roche; 3005372) which allows the quantification of caspase activity through the cleavage of a fluorescent substrate. DNA fragmentation has been estimated using the nuclear stain Hoescht 33345 (Sigma, B2261; and fluorescence microscopy to locate areas of chromatin condensation. Total viability of cell population has been quantified through measurement of the ability of mitochondrial reductase to metabolise the fluorescent substrate MTT (Sigma, M2128)(Levitz S.M. &
15 Diamond, R.D. J. Infect. Dis. 1985 Nov; 152(5):938-45).

Conditions for early, mid and late stage apoptosis in H9C2 and PHNC have been defined using hypoxia and nutrient-depleted growth medium to mimic those ischaemic conditions found *in vivo* (Brar, B.K., et al., J. Biol. Chem. 2000, 275, 8508-8514). Transduction of PHNC with EIAV vectors containing EGLN3 is sufficient to cause an increase in caspase activity in cells cultured under normoxic conditions,
20 confirming the role of EGLN3 in the induction of cardiomyocyte apoptosis. Using an MOI of 50, a 2-fold increase in caspase activity was seen in EGLN3 transduced cells, when compared to controls 48 hours post transduction (see Figure 32j).

Increased expression of EGL nine (*C.elegans*) homolog 3 in transduced cells is confirmed by TaqMan, immunofluorescence and Western blotting. Similar experiments are performed to determine whether EGL
25 nine (*C.elegans*) homolog 3 expression further sensitises H9C2 and PHNC to previously defined ischaemic insults. Staurosporine (Calbiochem; 569397) and Smart2.IRES.GFP EIAV vectors containing the Bax gene will be applied as chemical and viral pro-apoptotic controls, respectively (Yue, T-L., et al., J. Mol. Cell. Cardiol. 1998, 30, 495-507; Reed, J.C. J Cell Biol. 1994, 124(1-2):1-6).

30 Gene silencing approaches may be undertaken to down-regulate endogenous expression of EGLN3 in PHNC to determine the degree of protection against apoptotic cell death provided by a reduction in EGLN3 activity. RNA interference (RNAi) (Elbashir, SM et al., Nature 2001, 411, 494-498) is one method of sequence specific post-transcriptional gene silencing that may be employed. Short dsRNA oligonucleotides are synthesised *in vitro* and introduced into a cell. The sequence specific binding of

- these dsRNA oligonucleotides triggers the degradation of target mRNA, reducing or ablating target protein expression. A Hammerhead ribozyme library, contained in EIAV expression vectors, may also be applied. Efficacy of both gene silencing approaches may be assessed initially through the measurement of EGLN3 expression, at the RNA level by TaqMan and at the protein level by Western blotting. Protection
- 5 against previously described ischaemic insults provided by these methods of EGLN3 gene silencing may be assayed biologically as detailed above. Caspase inhibitors (caspase 3 inhibitor V, 2129002 and caspase inhibitor I, 627610, both Calbiochem) and Smart2.IRES.GFP EIAV vectors containing the Bcl-2 gene may be applied as chemical and viral anti-apoptotic controls, respectively (Kroemer, G. *Nat Med.* 1997, 3(6):614-20).
- 10 Similar "gain-of-function" and gene silencing approaches will be applied to the related gene, encoded by SEQ ID 90, named c1of12.

Genes with a greater response in renal epithelial cells

The dataset of Table 12 also contains genes which are induced preferentially in renal epithelial cells, in response to hypoxia. These genes are presented in Figures 39-44.

- 15 SeqID:117/118 EST
 SeqID:129/130 Hypothetical protein FLJ22622
 SeqID:31/32 TRIP-Br2
 SeqID:301/302 Tumor protein D52
 SeqID:91/92/92a Semaphorin 4b
- 20 SeqID:371/372 Dec-1

- For Semaphorin 4b (SeqID:91/92/92a), the clone presented in Figure 43 is p1P14, corresponding to IMAGE clone acc BE910319, the sequence of which covers a large region of the gene including protein coding sequence, which may cross-hybridise to other members of the semaphorin family. A separate clone (p1D17) as found in the original filing, was derived from the subtracted library and corresponds to a
- 25 more unique region of this gene in the untranslated region. From Table 12 it will be appreciated that a significant response is also found in the macrophage. This is validated by RNase protection assay data (see Figure 57). Further clarification of this gene using complementary experimentation methods will resolve the exact cell-type specific nature of the expression of this gene, though it is clear from this data that it is induced in renal epithelial cells and macrophages.

30 Genes with a greater response in mammary epithelial cells

The dataset of Table 12 also contains genes which are induced preferentially in mammary epithelial cells, in response to hypoxia. These genes are presented in Figures 45-52.

- SeqID:447/448 Calgranulin A
 SeqID:67/68 ERO1 (*S. cerevisiae*)-like
 SeqID:25/26 Hypothetical protein FLJ20500
 SeqID:229/230 N-myc downstream regulated
 5 SeqID:387/388 Decidual protein induced by progesterone
 SeqID:379/380 Integrin, alpha 5
 SeqID:225/226 Tissue factor
 SeqID:237/238 COX-2

In the case of Cox-2, which encodes a key drug target, it can be seen that in many cell types, especially
 10 the mammary epithelial cells, there is a clear induction in response to hypoxia. In contrast, for endothelial
 cells there is a very significant time-dependent *decrease* in Cox-2 gene expression in response to
 hypoxia. Similarly, for Calgranulin A, there is strong positive induction in hypoxic mammary epithelial
 cells, but in the macrophage, the response to hypoxia is negative. These clearly exemplify the unexpected
 finding that cell types respond to hypoxia differentially, both quantitatively but also qualitatively. This is
 15 not currently known.

Hypoxia regulation of Novel human genes

From Table 12, it will be appreciated that several genes with no prior annotation in public domain gene
 sequence databases are now identified as being regulated by hypoxia, in at least one cell type. To make
 this clear, these genes have been copied from Table 12 and presented in Tables 13 and 14), showing the
 20 hypoxia/ normoxia induction ratio of the cell type in which the response is most pronounced. These
 figures are derived by dividing the normalised expression value, as found in Table 5, in hypoxia by that in
 normoxia for the same cell type. In some cases, where hypoxia causes inhibition of gene expression, the
 fold change is prefixed by the term "DOWN". The cell type and time point of maximal response to
 hypoxia are also noted in Tables 13 and 14. The main purpose of Tables 13 and 14 is to demonstrate that
 25 these genes have significant responses to hypoxia *per se*.

In many cases, significant responses are seen in multiple cell types, though this data is not apparent here.
 In Table 13, the cDNA clones are currently un-annotated in public domain databases. In Table 14, the
 cDNA clones are currently annotated, but were not so as at the priority date.

Example 4: Additional disclosure of the effect of macrophage activation on hypoxia regulation of 30 gene expression

In Example 2, it is shown that activated and resting macrophages respond to hypoxia in different ways,
 showing that the hypoxia response is not a generic phenomenon. To consolidate this data, experiments
 were performed with the custom array, using additional experimental conditions and with a more in-depth

analysis. Significantly, the expression values used are not simple hypoxia/ normoxia ratios, done separately for macrophages of differing activation status, but rather the values used allow comparison of the relative expression levels throughout the entire set of experimental conditions. Hence, for any gene, all values throughout the entire set of experimental conditions are calculated by comparison to the median
 5 level of that gene throughout the dataset. This allows a clearer appreciation of the effects of hypoxia in the context of cell activation status. The following data demonstrates that of the newly discovered genes responsive to hypoxia, expression changes are also seen in response to key cytokines of the immune system, implying functions outside of the generic response to hypoxia and metabolism. This especially applies to unannotated genes, including ESTs and hypothetical proteins, showing potential functions in
 10 inflammation and angiogenesis on the basis of cytokine-regulation.

Macrophages were derived and cultured as described elsewhere in the specification. A total of 6 experimental conditions were analysed, as shown below. Where cells were treated with cytokines or hypoxia (0.1% oxygen), this was for 6 hr. Lipopolysaccharide (LPS) (from *E.coli* 026:B6; Sigma), gamma Interferon (IFN) and Interleukin-10 (IL-10) were all used at a final concentration of 100ng/ml.
 15 The effect of gamma Interferon and Lipopolysaccharide is to activate macrophages, with a Th1 biased phenotype, as found in many inflammatory conditions. Interleukin-10 is a Th2 cytokine and de-activates macrophages, and suppresses their effector functions.

Experimental

Condition

20	1.	No cytokines	Normoxia
	2.	No cytokines	Hypoxia
	3.	IL-10	Normoxia
	4.	IL-10	Hypoxia
	5.	LPS+IFN	Normoxia
25	6.	LPS+IFN	Hypoxia

In Table 15, genes are shown which respond to LPS+IFN in normoxia by producing at least a 2-fold increase in expression, indicating probable pro-inflammatory functions. From this dataset various patterns of hypoxia regulation will be appreciated on top of the effect of LPS+IFN.

For instance, the gene SCYA8 (p1121; SeqID: 479/480) is decreased in expression by hypoxia, changing
 30 from 0.54 to 0.18 between conditions #1 and #2. In condition #5 (LPS+IFN normoxia), expression is dramatically increased to a value of 19.6. When LPS+IFN is combined with hypoxia, this increase is dampened-down to a value of 12.2. So for this example, hypoxia and cell activation have opposing effects on gene expression. A similar expression profile is found for several other genes in Table 15.

In contrast, the gene P8 protein-candidate of metastasis 1 (p1F17; SeqID: 329/330) is increased in expression by hypoxia, changing from 0.26 to 1.78 between conditions #1 and #2. In condition #5 (LPS+IFN normoxia) expression is increased from condition #1 to a value of 1.16. In condition #6, (LPS+IFN normoxia) the expression is further increased to a value of 2.59. So for this example, hypoxia and cell activation have similar effects on expression (i.e. increases) and these are found to be synergistic. A similar expression profile is found for several other genes in Table 15, including for Semaphorin 4b (p1P14; SeqID:91/92/92a), which has been independently verified by RNase protection assay (see Figure 57).

A selection of novel genes taken from Table 15 is also presented as Figure 53. These novel genes are hence annotated here for the first time as being regulated not only by hypoxia, but also by Th1 inflammatory signals, as provided by LPS+IFN.

It will be appreciated that certain IMAGE clones were classed as novel and unannotated when the original patent filing was made (8 Dec 2000), but which can now be assigned to named genes. These are Uridine 5' monophosphate hydrolase 1 (clone p1I7; SeqID: 49/50) and Insulin induced protein 2 (clone p1D10; SeqID:75/76).

In Table 16, genes are shown which respond to LPS+IFN in normoxia by producing at least a 2-fold decrease in expression. From this dataset, various patterns of hypoxia regulation will be appreciated on top of the effect of LPS+IFN.

In Figure 54, novel genes from Table 16 which are down-regulated by LPS+IFN and up-regulated by hypoxia are presented. For most of these, the combined effect of LPS+IFN AND hypoxia produces only a minor induction above the level of expression for activated normoxic cells (for example p1F8/ SeqID:10/ Hypothetical Protein KIAA0914). In other cases, this is not the case, and hypoxia is able to over-ride the inhibitory effect of LPS+IFN on gene expression (for example p1D12/ SeqID:30/ Hypothetical Protein KIAA1376). This clearly demonstrates the finding that different cell types or physiological states of a cell type (as here), respond to hypoxia differently.

In Figure 55, novel genes from Table 16 which are down-regulated both by LPS+IFN and by hypoxia are presented. In many of the genes presented here, these stimuli are synergistic, with minimal expression obtained with a combination of LPS+IFN and hypoxia.

In Figure 56, a selection of named genes from Table 16 which are down-regulated by LPS+IFN, with various responses to hypoxia are presented. For the gene, Max-interacting Protein 1 two separate clones were available on the array corresponding to this gene (p1G5 from SeqID:280 and p1D22 from SeqID:120). In the original specification, the IMAGE clone corresponding to SeqID:120 (accession AA401496) was classified as an EST, and the IMAGE clone corresponding to SeqID:280 (accession

AA401496) was classified as "Max-interacting Protein 1", as determined by the UniGene database at that time. Now it is apparent that both of these clones correspond to Max-interacting Protein 1, explaining the similarity of their expression profiles in Figure 56. Clearly the response of this gene to hypoxia is inhibited by LPS+IFN.

- 5 The additional data showing effects of the Th1 activation stimulus LPS+IFN extends the finding of these genes as novel hypoxia regulated genes, and provides additional information about the relevance of these genes to disease mechanisms.

It will be appreciated that certain IMAGE clones were classed as novel and unannotated when the original patent filing was made (8 Dec 2000), but which can now be assigned to named genes. These are TRIP-

- 10 Br2 (clone p1D15; SeqID:31/32), MAX-interacting protein 1 (clone p1D22; SeqID:119/120).

In Tables 15 and 16 and Figures 53-56, showing genes which respond to LPS+IFN, it will be noticed that some of these genes also response to the inhibitory cytokine IL-10 (e.g. Semaphorin 4b, Hypothetical protein CGI-117). Other genes respond only to IL-10, but not to LPS+IFN. Specific responses to IL-10 are significant because this cytokine has been shown to have utility in suppressing inflammatory reactions

- 15 (Huizinga TW et al., *Rheumatology* 2000, 39: 1180-8).

Table 17 shows genes responsive to IL-10 (increased or decreased) but not affected significantly by LPS+IFN. Various patterns of hypoxia regulation will be appreciated.

Example 5: Gene expression in human tumors

- One of the utilities of the genes identified herein relates to the diagnosis and treatment of human tumors,
20 on the basis that hypoxia is frequently found in tumors.

- A study has been performed to examine the expression of these genes in a selection of breast and ovary tumors, comparing expression with normal adjacent tissue from the same patient. There is expected to be a large degree of variation between different patients, and the study here contains only 5 patients with a range of diagnoses. Therefore although certain genes will be identified from this data, other genes in the
25 current specification not flagged by this study are nevertheless likely to have utility in cancer.

Patients are designated as Letters:

E: 50 year old Caucasian female. Diagnosis: ovarian adenocarcinoma. Normal ovarian tissue derived from an age-matched separate individual.

- F: 60 year old female. Diagnosis: poorly differentiated adenocarcinoma. Normal ovarian tissue derived
30 from the same individual.

G: 41 year old female. Diagnosis: moderately-differentiated adenocarcinoma. Normal ovarian tissue derived from the same individual.

H: 40 year old female. Diagnosis: invasive ductal carcinoma. Normal breast tissue derived from the same individual.

- 5 K: 58 year old female. Diagnosis: invasive ductal carcinoma. Normal breast tissue derived from the same individual.

Data normalisation was done per-chip to correct for differences in labelling and hybridisation efficiency. Per-gene normalisation was done such that the expression values of each gene are relative to the median value of that gene throughout the series of samples. By comparing the expression values under normal
10 (nor) and tumor (tum) for a single patient, differences in expression between the normal and malignant tissue of that patient can be inferred.

In Table 18 are genes which are up-regulated at least 3-fold in at least one patient, comparing the tumor tissue to the adjacent normal tissue.

- 15 In Table 19 are genes which are down-regulated at least 3-fold in at least one patient, comparing the tumor tissue to the adjacent normal tissue.

Example 6: Effects of inflammatory cytokines on hypoxia-regulated genes

- Tumor necrosis factor alpha (TNF α) is a key pro-inflammatory cytokine both produced by and acting on the macrophage. The significance of TNF α to human disease is well established in the art. This is particularly the case in Rheumatoid arthritis and neutralising antibodies to TNF α have been
20 reported to offer clinical utility. Because hypoxia is another pathological condition exerted on macrophages in the synovia of RA patients, synergistic effects of these two stimuli are highly relevant to the discovery of novel inflammatory targets expressed by the macrophage. To investigate this, primary human macrophages were exposed to either hypoxia (0.1 % oxygen) or 100 ng/ml TNF α or to both for 6hr. The
25 data shown below provides further credence to the utility of the encoded proteins as inflammatory targets in macrophages and applies to *any* disease where hypoxia and TNF α are co-incident.

- Gene expression levels were measured and compared using the custom gene array. In data analysis per-gene normalisation was set up such that expression values represent the fold-change compared with the expression in untreated normoxic cells. Genes which are increased in expression in response to TNF α by
30 at least 2-fold, in either normoxic or hypoxic cells, are shown in Table 20. Genes which are decreased in expression in response to TNF α by at least 2-fold, in either normoxic or hypoxic cells, are shown in Table 21.

Another inflammatory cytokine implicated in diseases where hypoxia is frequently found is Interleukin-17 (IL-17). For example, this cytokine has been shown to mediate inflammation and joint destruction in arthritis (Lubberts et al *J Immunol* 2001 167:1004-1013). IL-17 has also been shown to stimulate macrophages to release other key pro-inflammatory cytokines (Jovanovic et al *J Immunol* 1998 160:3513-21). Therefore genes which respond to both hypoxia and IL-17 are especially likely to be relevant to disease processes and have utility in the design of therapeutic products. Genes which are increased in expression in response to IL-17 by at least 2-fold, in either normoxic or hypoxic cells, are shown in Table 22. Genes which are decreased in expression in response to IL-17 by at least 2-fold, in either normoxic or hypoxic cells, are shown in Table 23.

The cytokine IL-15 is implicated in several disease in which macrophages and hypoxia both feature as elements of the inflammatory state, such as in atherosclerosis (Wuttge DM et al *Am J Pathol.* 2001 159:417-23) and rheumatoid arthritis (McInnes IB et al *Immunol Today.* 1998 19:75-9). Although the main target of IL-15 is T-cells effects have also been shown on monocytes (Badolato R et al *Blood.* 1997 90:2804-9). Therefore genes which respond to both hypoxia and IL-15 are especially likely to be relevant to disease processes and have utility in the design of therapeutic products. Genes which are increased in expression in response to IL-15 by at least 2-fold, in either normoxic or hypoxic cells, are shown in Table 24. Genes which are decreased in expression in response to IL-15 by at least 2-fold, in either normoxic or hypoxic cells, are shown in Table 25.

Example 7: Rat foetal cardiomyocytes

Primary rat foetal cardiomyocytes provide an attractive experimental model for studying the responses of cardiac cells to ischaemia. Cells are obtained which are non-immortalised and which are seen to contract or beat in culture. It is of interest to examine how the responses of these cells to hypoxia (or related experimental conditions) compared and contrasts to other cell types. These other cell types might include those that are similarly sensitive to the effects of hypoxia (such as neurones) or might be cells that show a higher tolerance to hypoxia (such as macrophages). Experiments are performed in parallel for cardiomyocytes and other cell type(s). The responses of these specific cell types is then determined by hybridising labelled mRNA to microarrays. Alternative methods will include the construction of subtracted cDNA libraries for the individual treated cell types and assessing which genes are contained therein by sequencing.

Methods

Cardiomyocytes are harvested from heart ventricles of embryos aged E18 days, using a cell isolation kit (Neonatal cardiomyocyte isolation system; Worthington Biochemical Corporation, Lakewood, New

Jersey, 08701). They are seeded at 5×10^6 cells/100cm diameter petri dish in DMEM/M199, 10% horse serum, 5% FCS, 1% penicillin, streptomycin, glutamine for 5 days at 37C. Media is changed during the 5 days.

Other cell types used for comparison with cardiomyocytes, are cultured according to their optimum conditions and/ or the standard routine. These cell types may include cardiomyocytes in a different physiological setting, such as in an intact beating heart, or a different developmental state of the cardiomyocyte, such as cardiomyoblast.

Identical seeded petri dishes are placed either in a standard tissue culture incubator (95% air/ 5% CO₂) or in a hypoxia incubator (0.1% oxygen / 5% CO₂ / 0.1% oxygen for 6 hours. This is done separately for both cardiomyocytes and the other cell type(s) to be compared. Other experimental conditions might more closely approximate ischemia, by incorporating components additional to hypoxia.

At the end of the exposure to hypoxia, cells are placed on a chilled platform, washed in cold PBS and total RNA is extracted using RNazol B (Tel-Test, Inc; distributed by Biogenesis Ltd) following the manufacturers instructions. Where appropriate, polyadenylated mRNA is extracted from the total RNA using a commercial kit following the manufacturers instructions (Promega; PolyA Tract mRNA isolation System IV).

Array hybridisations and construction/analysis of subtracted cDNA libraries are performed according to standard methods or as described elsewhere in this specification.

Example 8: Comparison of the hypoxic-responses between populations of rat primary cultured neurons by a subtraction cloning / array screening approach.

Different regions of the central nervous system display different sensitivities to hypoxia and to ischaemia. Susceptibility to tissue damage in this manner may occur as a result of intrinsic differences in gene expression between cells. To evaluate this hypothesis, primary cultures of rat neurons from different regions of the brain are established. Cultures are exposed to various experimental conditions which are pertinent to pathologies of the hypoxic/ischemic brain. These would include hypoxic insults as have been described, or to hypoxia/ischaemia where the conditions more closely approximate pathological ischemia. Either condition may be preceded by prior hypoxic-preconditioning, where transient exposure to hypoxia renders cells less sensitive to subsequent acute treatment. For all possible experimental treatments, a similar routine is performed for distinct neuron subtypes, in order to compare their responses. Such comparisons may be made by hybridizing labelled mRNA to microarrays or derivatives thereof. Alternatively subtracted libraries might be constructed individually for each treated neuron subtype, and clones which are confirmed to be changed in expression to be sequenced. The collection of genes arising from the different neuron subtypes will be compared.

Methods

Primary cultures are established according to standard procedures from embryonic rats aged from E14 to E18 (Dunnett SB, Bjorkland A (Eds.) 1992. *Neural Transplantation, A Practical Approach*. IRL Press). Isolated neurons include but are not limited to those from ventral mesencephalon, striatum, hippocampus, cerebellum, cerebral cortex, dorsal root ganglia and superior cervical ganglia.

Cells are maintained in culture for 3-14 days in humidified culture incubators at 37°C, 5% CO₂, 95% air (Normoxia) in Neurobasal Medium (Brewer GJ, 1995, *Journal of Neuroscience Research* 42:674-83) supplemented with B27 (both Life Technologies). For the hypoxia-preconditioning, cells are transferred to a second incubator at 37°C, 5% CO₂, 94.9% Nitrogen, 0.1% Oxygen (Hypoxia) for 30-180 minutes and returned to the normoxic incubator for 24 hours (Pringle *et al.*, 1997, *Neuropathology and Applied Neurobiology* 23:289-298). For the hypoxic stimulus, either independent from or subsequent to hypoxia-preconditioning, cells are transferred to the hypoxic incubator for 2-6 hours as determined in time course experiments. Additionally, as appropriate, the medium in which the cells are grown is replaced with glucose-free media for establishment of experimental ischaemia (Ray AM, Owen DE, Evans ML, Davis JB Benham, 2000. Caspase inhibitors are functionally neuroprotective against oxygen glucose deprivation induced CA1 death in rat organotypic hippocampal slices). At the end of the exposure to hypoxia (or hypoxia/ischaemia), cells are, placed on a chilled platform, washed in cold PBS and total RNA is extracted using RNazol B (Tel-Test, Inc; distributed by Biogenesis Ltd) following the manufacturers instructions. Where appropriate, polyadenylated mRNA is extracted from the total RNA using a commercial kit following the manufacturers instructions (Promega; PolyA Tract mRNA isolation System IV).

Array hybridisations and construction/analysis of subtracted cDNA libraries are performed according to standard methods or as described elsewhere in this specification.

Example 9: Semaphorin 4b

We have screened cDNA libraries derived from the human brain and leukocytes, to obtain an unequivocal and accurate full length cDNA sequence (SEQ ID No 92a) and the accurate presumptive amino acid sequence (SEQ ID No 91).

The amino acid sequence above was derived by taking the first ATG. We have various independent lines of evidence that this is the *bona fide* translation initiation codon.

Basic analysis of this sequence, reveals the following motifs:

signal peptide (pSORT) Start: 1 End: 37;

Transmembrane (pSORT) Start: 718 End: 734;

cleavage site (pSORT) Start: 38 End: 38;

- Proline rich region Start: 758 End: 824;
Sema domain (pfam) Start: 70 End: 503;
Plexin repeat (pfam) Start: 525 End: 548;
integrin, beta domain (pfam) Start: 532 End: 546;
5 cytoplasmic tail Start: 735 End: 837.

To confirm the hypoxic regulation of Sema4b, we used RNase protection assay (see Figure 57). Hypoxia is a feature of several inflammatory conditions often accompanied by superoxide radicals and the immune regulator gamma interferon. In this experiment we have made the following findings:

- Expression is activated by hypoxia (3.3 fold)
- 10 • Expression is activated by gamma interferon and LPS (3.9 fold)
- Expression is activated synergistically by hypoxia plus gamma interferon/ LPS (7.3 fold)
- Expression is activated by superoxide radicals (5.0 fold)

To investigate the size of the mRNA and the tissue distribution, Northern blotting was done (see Figure 58). This shows that the gene is expressed as a single transcript at relatively low levels in unstimulated
15 human tissues.

We have also found that a molecule that is probably associated with Semaphorin 4B, called psd-95 is another macrophage hypoxia-induced protein (see SEQ ID No 299). This is based on the fact that psd-95 binds the cytoplasmic tail of Sema4c (Inagaki et al., J Biol Chem. 2001; 276(12): 9174-81), which like Sema4b, contains proline rich sequence. Therefore, both Semaphorin 4B, and a probable partner are co-
20 ordinatorily regulated by hypoxia.

Example 10: Discussion of relevance of individual clones

The Oxford BioMedica clone p1F12 represents Hypothetical protein FLJ13611. The protein sequence encoded by Hypothetical protein FLJ13611 is represented in the public databases by the accession NP_079217 and is described in this patent by Seq ID 1. The nucleotide sequence is represented in the
25 public sequence databases by the accession NM_024941 and is described in this patent by Seq ID 2. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1F2 represents Hypothetical protein FLJ20037. The protein sequence
30 encoded by Hypothetical protein FLJ20037 is represented in the public databases by the accession CAB65981 and is described in this patent by Seq ID 3. The nucleotide sequence is represented in the public sequence databases by the accession NM_017633 and is described in this patent by Seq ID 4. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore

implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Hypothetical protein FLJ20037 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient. We expect
5 increased activity of the gene product to have an anti-tumour effect.

The Oxford BioMedica clone p1F10 represents Hypothetical protein DKFZp434P0116. The protein sequence encoded by Hypothetical protein DKFZp434P0116 is represented in the public databases by the accession T46364 and is described in this patent by Seq ID 5. The nucleotide sequence is represented in the public sequence databases by the accession NM_017593 and is described in this patent by Seq ID 6.
10 Hypothetical protein DKFZp434P0116 is predicted to be a kinase due to high structural similarity with other known kinases. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypothetical protein DKFZp434P0116 is repressed in macrophages activated by LPS and gamma interferon. We expect it to have an anti-inflammatory role.

15 The Oxford BioMedica clone p1F19 represents Hypothetical protein KIAA0212. The protein sequence encoded by Hypothetical protein KIAA0212 is represented in the public databases by the accession BAA13203 and is described in this patent by Seq ID 7. The nucleotide sequence is represented in the public sequence databases by the accession NM_014674 and is described in this patent by Seq ID 8. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore
20 implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1F8 represents Hypothetical protein KIAA0914. The protein sequence encoded by Hypothetical protein KIAA0914 is represented in the public databases by the accession NP_055698 and is described in this patent by Seq ID 9. The nucleotide sequence is represented in the
25 public sequence databases by the accession NM_014883 and is described in this patent by Seq ID 10. Hypothetical protein KIAA0914 shows high structural similarity to Human Class I alpha 1,2-Mannosidase and conservation of active site and binding site residues, therefore we predict that Hypothetical protein KIAA0914 will act as a mannosidase. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have
30 utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Hypothetical protein

KIAA0914 is repressed in macrophages activated by LPS and gamma interferon. We expect the gene product to have an anti-inflammatory role.

The Oxford BioMedica clone p1F5 represents Hypothetical protein FLJ20281. The protein sequence encoded by Hypothetical protein FLJ20281 is represented in the public databases by the accession
5 XP_008736 and is described in this patent by Seq ID 11. The nucleotide sequence is represented in the public sequence databases by the accession NM_017742 and is described in this patent by Seq ID 12. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be
10 central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Hypothetical protein FLJ20281 is induced in macrophages activated by TNFalpha.

The Oxford BioMedica clone p1F18 represents Hypothetical protein KIAA0876. The protein sequence
15 encoded by Hypothetical protein KIAA0876 is represented in the public databases by the accession BAA74899 and is described in this patent by Seq ID 13. The nucleotide sequence is represented in the public sequence databases by the accession XM_035625 and is described in this patent by Seq ID 14. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic
20 products.

The Oxford BioMedica clone p1F7 represents Spectrin, beta, non-erythrocytic 1. The protein sequence encoded by Spectrin, beta, non-erythrocytic 1 is represented in the public databases by the accession NP_003119 and is described in this patent by Seq ID 15. The nucleotide sequence is represented in the public sequence databases by the accession NM_003128 and is described in this patent by Seq ID 16.
25 Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Spectrin, beta, non-erythrocytic 1 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient. We expect
30 increased activity of the gene product to have an anti-tumour effect.

The Oxford BioMedica clone p1F21 represents Hematopoietic Zinc finger protein. The protein sequence encoded by Hematopoietic Zinc finger protein is represented in the public databases by the accession AAL08625 and is described in this patent by Seq ID 17. The nucleotide sequence is represented in the

public sequence databases by the accession AK024404 and is described in this patent by Seq ID 18. Hematopoietic Zinc finger protein is a transcriptional regulator that contains a Cys2-His2 zinc finger motif. It is predicted to bind to metal response elements (MRE) and therefore activate the transcription of genes that contain a MRE sequence within their promoter region such as metallothioneins. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD. Hematopoietic Zinc finger protein is preferentially induced by hypoxia in monocytes or macrophages and a restricted number of other cell types. It is therefore a candidate for specific intervention for treatment or diagnosis of the above diseases.

The Oxford BioMedica clone p1F9 represents Hypothetical protein KIAA0742. The protein sequence encoded by Hypothetical protein KIAA0742 is represented in the public databases by the accession NP_060903 and is described in this patent by Seq ID 19. The nucleotide sequence is represented in the public sequence databases by the accession AB018285 and is described in this patent by Seq ID 20. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Hypothetical protein KIAA0742 is repressed in macrophages activated by LPS and gamma interferon. We expect it to have an anti-inflammatory role. Hypothetical protein KIAA0742 shows significant homology to the transcription factor hairless. We therefore propose that Hypothetical protein KIAA0742 may play a crucial role in the regulation of hair growth. Accordingly, this aspect of the invention includes the use of this protein, fragments and functional equivalents of this protein, encoding nucleic acid molecules, in addition to ligands that bind specifically to this protein, in the diagnosis and treatment of hair loss.

The Oxford BioMedica clone p1E13 represents Hypothetical protein PRO0823. The protein sequence encoded by Hypothetical protein PRO0823 is represented in the public databases by the accession AAF71073 and is described in this patent by Seq ID 21. The nucleotide sequence is represented in the public sequence databases by the accession AF116653 and is described in this patent by Seq ID 22. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore

implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Hypothetical protein PRO0823 is repressed in macrophages activated by LPS and gamma interferon. We expect it to have an anti-inflammatory role. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Hypothetical protein PRO0823 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient. We expect increased activity of the gene product to have an anti-tumour effect.

The Oxford BioMedica clones p1D1 and p1D2 represent the Hypothetical protein FLJ10134. The protein sequence encoded by Hypothetical protein FLJ10134 is represented in the public databases by the accession NP_060474 and is described in this patent by Seq ID 23. The nucleotide sequence is represented in the public sequence databases by the accession NM_018004 and is described in this patent by Seq ID 24. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. HIF1alpha and EPAS1 are transcription factors that mediate the response to hypoxia of several genes, and have themselves been implicated in specific diseases. By adenoviral over-expression of EPAS1 we show augmentation of the hypoxic induction of certain genes, further confirming their status as responsive to hypoxia. Hypothetical protein FLJ10134 has been shown to be induced by hypoxia to a greater degree following adenoviral over-expression of EPAS1. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Hypothetical protein FLJ10134 is repressed in macrophages activated by LPS and gamma interferon. We expect it to have an anti-inflammatory role. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Hypothetical protein FLJ10134 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient. We expect increased activity of the gene product to have an anti-tumour effect.

The Oxford BioMedica clone p1D4 represents Hypothetical protein FLJ20500. The protein sequence encoded by Hypothetical protein FLJ20500 is represented in the public databases by the accession NP_061931 and is described in this patent by Seq ID 25. The nucleotide sequence is represented in the public sequence databases by the accession NM_019058 and is described in this patent by Seq ID 26. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore

implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypothetical protein FLJ20500 is preferentially induced by hypoxia in mammary epithelial cells.

The Oxford BioMedica clone p1D9 represents Hypothetical protein DKFZP564D116. The protein
5 sequence encoded by Hypothetical protein DKFZP564D116 is represented in the public databases by the accession T08708 and is described in this patent by Seq ID 27. The nucleotide sequence is represented in the public sequence databases by the accession AL050022 and is described in this patent by Seq ID 28. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic
10 products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Hypothetical protein DKFZP564D116 is repressed in macrophages activated by LPS and gamma interferon. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been
15 shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Hypothetical protein DKFZP564D116 is induced in macrophages activated by TNFalpha.

The Oxford BioMedica clone p1D12 represents Hypothetical protein KIAA1376. The protein sequence
20 encoded by Hypothetical protein KIAA1376 is represented in the public databases by the accession BAA92614 and is described in this patent by Seq ID 29. The nucleotide sequence is represented in the public sequence databases by the accession AB037797 and is described in this patent by Seq ID 30. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic
25 products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Hypothetical protein KIAA1376 is repressed in macrophages activated by LPS and gamma interferon. We expect it to have an anti-inflammatory role.

30 The Oxford BioMedica clone p1D15 represents TRIP-Br2. The protein sequence encoded by TRIP-Br2 is represented in the public databases by the accession NP_055570 and is described in this patent by Seq ID 31. The nucleotide sequence is represented in the public sequence databases by the accession NM_014755 and is described in this patent by Seq ID 32. TRIP-BR2 is a PHD zinc finger and bromodomain interacting protein transcriptional regulator and is involved in the regulation of cell cycle

progression. Its hypoxia-regulation is likely to have important disease-relevant effects. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. The response of renal epithelial cells to hypoxia is pertinent to kidney failure, especially regarding the
5 medullary tissue. TRIP-Br2 is preferentially induced by hypoxia in renal epithelial cells. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. TRIP-Br2 is repressed in macrophages activated by LPS and gamma interferon.

- 10 The Oxford BioMedica clone p1D16 represents Hypothetical protein FLJ20308. The protein sequence encoded by Hypothetical protein FLJ20308 is represented in the public databases by the accession XP_039852 and is described in this patent by Seq ID 33. The nucleotide sequence is represented in the public sequence databases by the accession AK000315 and is described in this patent by Seq ID 34. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore
15 implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Hypothetical protein FLJ20308 is repressed in macrophages activated by LPS and gamma
20 interferon. We expect it to have an anti-inflammatory role.

The Oxford BioMedica clone p1J13 represents Hypothetical nuclear factor SBB122. The protein sequence encoded by Hypothetical nuclear factor SBB122 is represented in the public databases by the accession NP_065128 and is described in this patent by Seq ID 35. The nucleotide sequence is represented in the public sequence databases by the accession NM_020395 and is described in this patent by Seq ID 36.
25 Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1I22 represents Hypothetical protein KIAA1429. The protein sequence encoded by Hypothetical protein KIAA1429 is represented in the public databases by the accession
30 BAA92667 and is described in this patent by Seq ID 37. The nucleotide sequence is represented in the public sequence databases by the accession AB037850 and is described in this patent by Seq ID 38. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1J6 represents Hypothetical protein FLJ10206. The protein sequence encoded by Hypothetical protein FLJ10206 is represented in the public databases by the accession AAH06108 and is described in this patent by Seq ID 39. The nucleotide sequence is represented in the public sequence databases by the accession NM_018025 and is described in this patent by Seq ID 40.

- 5 Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially
10 relevant. Hypothetical protein FLJ10206 is repressed in macrophages activated by IL-17 and is also repressed in macrophages activated by IL-15. These are pro-inflammatory cytokines, and we expect the hypothetical protein FLJ10206 to have an anti-inflammatory role.

- The Oxford BioMedica clone p1I5 represents Hypothetical protein FLJ10815. The protein sequence encoded by Hypothetical protein FLJ10815 is represented in the public databases by the accession
15 BAA91830 and is described in this patent by Seq ID 41. The nucleotide sequence is represented in the public sequence databases by the accession NM_018231 and is described in this patent by Seq ID 42. Hypothetical protein FLJ10815 is structurally similar to an alpha / beta barrel structure. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.
20 Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Hypothetical protein FLJ10815 is repressed in macrophages activated by LPS and gamma interferon. We expect it to have an anti-inflammatory role.

- 25 The Oxford BioMedica clone p1I13 represents Hypothetical protein FLJ11100. The protein sequence encoded by Hypothetical protein FLJ11100 is represented in the public databases by the accession NP_060701 and is described in this patent by Seq ID 43. The nucleotide sequence is represented in the public sequence databases by the accession NM_018321 and is described in this patent by Seq ID 44. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore
30 implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1I17 represents Hypothetical protein FLJ20644. The protein sequence encoded by Hypothetical protein FLJ20644 is represented in the public databases by the accession NP_060387 and is described in this patent by Seq ID 45. Hypothetical protein FLJ20644 is a putative

Serine/threonine phosphatase. Region 250 – 450 shows high structural similarity to other Serine/threonine phosphatases. The nucleotide sequence is represented in the public sequence databases by the accession NM_017917 and is described in this patent by Seq ID 46. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have
5 utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p115 represents Hypothetical protein CGI-117. The protein sequence encoded by Hypothetical protein CGI-117 is represented in the public databases by the accession Q9Y3C1 and is described in this patent by Seq ID 47. The nucleotide sequence is represented in the public sequence databases by the accession NM_016391 and is described in this patent by Seq ID 48.
10 Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. HIF1alpha and EPAS1 are transcription factors that mediate the response to hypoxia of several genes, and have themselves been implicated in specific diseases. By adenoviral over-expression of HIF1alpha or EPAS1 we show augmentation of the hypoxic induction of certain genes, further
15 confirming their status as responsive to hypoxia. Hypothetical protein CGI-117 has been shown to be induced by hypoxia to a greater degree following adenoviral over-expression of either HIF1alpha or EPAS1. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially
20 relevant. Hypothetical protein CGI-117 is repressed in macrophages activated by LPS and gamma interferon. We expect it to have an anti-inflammatory role.

The Oxford BioMedica clone p117 represents Uridine 5' monophosphate hydrolase 1. The protein sequence encoded by Uridine 5' monophosphate hydrolase 1 is represented in the public databases by the accession NP_057573 and is described in this patent by Seq ID 49. The nucleotide sequence is
25 represented in the public sequence databases by the accession NM_016489 and is described in this patent by Seq ID 50. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been
30 shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Uridine 5' monophosphate hydrolase 1 is induced in macrophages activated by LPS and gamma interferon and is also induced in macrophages activated by IL-15. We expect it to have a pro-inflammatory role, and its inhibition may have an anti-inflammatory effect.

The protein sequence encoded by Hypothetical protein KIAA0014 is represented in the public databases by the accession NP_055480 and is described in this patent by Seq ID 51. The nucleotide sequence is represented in the public sequence databases by the accession NM_014665 and is described in this patent by Seq ID 52. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p114 represents Hypothetical protein HSPC196. The protein sequence encoded by Hypothetical protein HSPC196 is represented in the public databases by the accession NP_057548 and is described in this patent by Seq ID 53. The nucleotide sequence is represented in the public sequence databases by the accession NM_016464 and is described in this patent by Seq ID 54. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Hypothetical protein HSPC196 is repressed in macrophages activated by LPS and gamma interferon. We expect it to have an anti-inflammatory role.

The Oxford BioMedica clone p118 represents Hypothetical protein FLJ11296. The protein sequence encoded by Hypothetical protein FLJ11296 is represented in the public databases by the accession XP_004747 and is described in this patent by Seq ID 55. The nucleotide sequence is represented in the public sequence databases by the accession NM_018384 and is described in this patent by Seq ID 56. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1116 represents Hypothetical protein KIAA1668. The protein sequence encoded by Hypothetical protein KIAA1668 is represented in the public databases by the accession BAB33338 and is described in this patent by Seq ID 57. The nucleotide sequence is represented in the public sequence databases by the accession AB051455 and is described in this patent by Seq ID 58. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1111 represents SECIS binding protein 2. The protein sequence encoded by SECIS binding protein 2 is represented in the public databases by the accession AAK57518 and is

described in this patent by Seq ID 59. The nucleotide sequence is represented in the public sequence databases by the accession AF380995 and is described in this patent by Seq ID 60. SECIS binding protein 2 is a crucial component in the complex required for the translation of mammalian selenoprotein mRNAs. Selenoproteins are important responders to redox conditions and many selenoproteins are known to protect from cell death. Our demonstration of the hypoxia induction of SECIS binding protein 2 opens new avenues for diagnosis and therapeutic intervention. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, SECIS binding protein 2 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient. We expect increased activity of the gene product to have an anti-tumour effect.

The Oxford BioMedica clone p1E8 represents cDNA: FLJ22249 fis, clone HRC02674. The sequence cDNA: FLJ22249 fis, clone HRC02674 is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AK025902 and is described in this patent by Seq ID 62. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1E18 represents Plexin C1. The protein sequence encoded by Plexin C1 is represented in the public databases by the accession NP_005752 and is described in this patent by Seq ID 63. The nucleotide sequence is represented in the public sequence databases by the accession NM_005761 and is described in this patent by Seq ID 64. Plexins are a large family of receptors for transmembrane, secreted, and GPI-anchored semaphorins in vertebrates and play a significant role in signal transduction [Tamagnone et al 1999, Cell 99:71-80]. Elsewhere in this patent we disclose hypoxic regulation of a new semaphorin 4b, and we propose co-regulation of these molecules by hypoxia and their relevance to inflammatory disease, its diagnosis and therapy. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Plexin C1 is repressed in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1E16 represents cDNA DKFZp586E1624. The sequence cDNA DKFZp586E1624 is not represented in the public databases by a protein accession. The nucleotide

sequence is represented in the public sequence databases by the accession AL110152 and is described in this patent by Seq ID 66. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. HIF1alpha and EPAS1 are transcription factors that mediate the response to hypoxia of several genes, and have themselves been implicated in specific diseases. By adenoviral over-expression of EPAS1 we show augmentation of the hypoxic induction of certain genes, further confirming their status as responsive to hypoxia. Its preferential regulation by EPAS1 provides a route to preferential intervention, to avoid toxicity to other tissues. The cDNA DKFZp586E1624 has been shown to be induced by hypoxia to a greater degree following adenoviral over-expression of EPAS1.

10 Endothelial cells are key to angiogenesis, a process implicated in several diseases associated with hypoxia, including cancer and rheumatoid arthritis. The cDNA DKFZp586E1624 is preferentially induced by hypoxia in endothelial cells. We expect this gene product to have a pro-angiogenic effect, and its inhibition to have an anti-angiogenic effect.

The Oxford BioMedica clones p1D5 and p1D6 represent ERO1 (*S. cerevisiae*)-like. The protein sequence encoded by ERO1 (*S. cerevisiae*)-like is represented in the public databases by the accession NP_055399 and is described in this patent by Seq ID 67. The nucleotide sequence is represented in the public sequence databases by the accession NM_014584 and is described in this patent by Seq ID 68. ERO1 (*S. cerevisiae*)-like has been shown to be a flavin adenine dinucleotide (FAD) binding protein. Binding of FAD enables ERO1 (*S. cerevisiae*)-like to oxidise protein disulfide isomerase (PDI). We propose that the oxidation of PDI by ERO1 (*S. cerevisiae*)-like stops PDI autodegradation, therefore increasing levels of the protein. Increased levels of PDI have been shown to be neuroprotective by inhibiting apoptotic cell death. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. HIF1alpha and EPAS1 are transcription factors that mediate the response to hypoxia of several genes, and have themselves been implicated in specific diseases. By adenoviral over-expression of EPAS1 we show augmentation of the hypoxic induction of certain genes, further confirming their status as responsive to hypoxia. ERO1 (*S. cerevisiae*)-like has been shown to be induced by hypoxia to a greater degree following adenoviral over-expression of EPAS1. Its preferential regulation by EPAS1 provides a route to preferential intervention, to avoid toxicity to other tissues. ERO1 (*S. cerevisiae*)-like is preferentially induced by hypoxia in mammary epithelial cells. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. ERO1 (*S. cerevisiae*)-like is repressed in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human

tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, ERO1 (*S. cerevisiae*)-like is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1E12 represents Hypothetical protein DKFZP434E1723. The protein
5 sequence encoded by Hypothetical protein DKFZP434E1723 is represented in the public databases by the accession XP_05338 and is described in this patent by Seq ID 69. The nucleotide sequence is represented in the public sequence databases by the accession BC010005 and is described in this patent by Seq ID 70. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic
10 products.

The Oxford BioMedica clone p1E10 represents cDNA FLJ11041 fis clone PLACE1004405. The sequence encoded by cDNA FLJ11041 fis, clone PLACE1004405 is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AK001903 and is described in this patent by Seq ID 72. Hypoxia is an important feature
15 of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites; so gene expression responses to both hypoxia and cytokines are especially relevant. The cDNA FLJ11041 fis
20 clone PLACE1004405 is induced in macrophages activated by LPS and gamma interferon. We expect it to have a pro-inflammatory role, and its inhibition may have an anti-inflammatory effect.

The Oxford BioMedica clone p1C21 represents Tubulin, beta, 4. The protein sequence encoded by Tubulin, beta, 4 is represented in the public databases by the accession NP_006077 and is described in this patent by Seq ID 73. The nucleotide sequence is represented in the public sequence databases by the
25 accession NM_006086 and is described in this patent by Seq ID 74. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1D10 represents Insulin induced protein 2. The protein sequence encoded by Insulin induced protein 2 is represented in the public databases by the accession AAD43048 and is
30 described in this patent by Seq ID 75. The nucleotide sequence is represented in the public sequence databases by the accession AF125392 and is described in this patent by Seq ID 76. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Insulin induced protein 2 is induced in macrophages activated by LPS and gamma interferon. We expect it to have a pro-inflammatory role, and its inhibition may have an anti-inflammatory effect.

The Oxford BioMedica clones p1D13 and p1A22 represent Adenylate kinase 3. The protein sequence encoded by Adenylate kinase 3 is represented in the public databases by the accession NP_037542 and is described in this patent by Seq ID 77 and 263. The nucleotide sequence is represented in the public sequence databases by the accession NM_013410 and is described in this patent by Seq ID 78 and 264.

10 Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Adenylate kinase 3 is induced in macrophages activated by LPS and gamma interferon. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Adenylate kinase 3 is induced in macrophages
15 activated by TNFalpha.

The Oxford BioMedica clone p1E9 represents a novel PI-3-kinase adapter. The protein sequence encoded by the novel PI-3-kinase adapter is not represented in the public databases by a protein accession but is described in this patent by Seq ID 79. The nucleotide sequence of an unannotated EST corresponding to the novel PI-3-kinase adapter is represented in the public sequence databases by the accession R62339
25 and is described in this patent by Seq ID 80. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products
30 for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD and peripheral arterial disease. The novel PI-3-kinase adapter is preferentially induced by hypoxia in monocytes or macrophages, indicating utility of the encoded protein in the design of therapeutic, prognostic and diagnostic products addressing diseases involving macrophages and hypoxia. In a gene array analysis it is

expressed in hypoxic monocytes and macrophages at levels 6-fold higher than the median expression level of this gene throughout 9 other cell types in either normoxia or hypoxia. In more sensitive TaqMan analysis the novel PI-3-kinase adapter it is found to be expressed at approximately 1000 times the levels of 9 other cell types, all exposed to hypoxia for 18hr. The relevance of the novel PI-3-kinase adapter to human disease is also appreciated from comparison with a related murine gene, BCAP. It is known that this gene is phosphorylated by the tyrosine kinase, Syk. We also show novel data regarding Syk, in that it is also induced in response to hypoxia in a tissue specific manner identical to that of the novel PI-3-kinase adapter. Therefore the biological relevance and utility of our discovery of hypoxic induction of the novel PI-3-kinase adapter gene is further highlighted.

10

The Oxford BioMedica clone p1F1 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA489477 and is described in this patent by Seq ID 82. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1E7 represents a novel Metallothionein. The protein sequence encoded by Novel Metallothionein is not represented in the public databases by a protein accession but is described in this patent by Seq ID 83. The nucleotide sequence is represented in the public sequence databases by the accession R06601 and is described in this patent by Seq ID 84. Metallothioneins can act as an antioxidant and free-radical scavenger and are therefore protective against cell death in hypoxia. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. HIF1alpha and EPAS1 are transcription factors that mediate the response to hypoxia of several genes, and have them selves been implicated in specific diseases. By adenoviral over-expression of HIF1alpha we show augmentation of the hypoxic induction of certain genes, further confirming their status as responsive to hypoxia. The novel Metallothionein represented by Seq ID 84 has been shown to be induced by hypoxia to a greater degree following adenoviral over-expression of HIF1alpha. Hepatocytes are the main cell type of the liver and genes that are induced in response to hypoxia in this cell type are relevant to development of diagnostics and therapeutics towards liver diseases involving hypoxia, including cirrhosis. The novel Metallothionein represented by Seq ID 84 is preferentially induced by hypoxia in hepatocytes. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been

shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. The novel Metallothionein represented by Seq ID 84 is induced in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1E6 represents EGL nine (C.elegans) homolog 3. The protein sequence encoded by EGL nine (C.elegans) homolog 3 is represented in the public databases by the accession NP_071356 and is described in this patent by Seq ID 85. The nucleotide sequence is represented in the public sequence databases by the accession NM_022073 and is described in this patent by Seq ID 86. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. HIF1alpha and EPAS1 are transcription factors that mediate the response to hypoxia of several genes, and have themselves been implicated in specific diseases. By adenoviral over-expression of EPAS1 we show augmentation of the hypoxic induction of certain genes, further confirming their status as responsive to hypoxia. EGL nine (C.elegans) homolog 3 has been shown to be induced by hypoxia to a greater degree following adenoviral over-expression of EPAS1. Its preferential regulation by EPAS1 provides a route to preferential intervention, to avoid toxicity to other tissues. Hepatocytes are the main cell type of the liver and genes that are induced in response to hypoxia in this cell type are relevant to development of diagnostics and therapeutics towards liver diseases involving hypoxia, including cirrhosis. EGL nine (C.elegans) homolog 3 is preferentially induced by hypoxia in hepatocytes. We find that EGLN3 and a related human gene Clorf12 (seq ID 89/90) both of which are predicted to be proline hydroxylases, are expressed at differing absolute expression levels in different tissues. For instance, in the hypoxic hepatocyte (6hr) the normalised expression values of EGLN and clorf12 are 0.015 and 0.0074 respectively, i.e. EGLN being the dominant gene. In contrast, in the neuroblastoma cell line SH-SY5Y, the normalised expression values of EGLN and clorf12 after 6hr hypoxia are 0.0012 and 0.108 respectively, i.e. clorf12 being the dominant gene by a large margin. This data demonstrates that c1ORF12 and EGLN3 are not constitutively expressed at an equal amount in different tissues indicating specificity of function. Therefore therapeutic products may be developed based on this data, with the goal of modulating proline hydroxylation of target proteins (such as HIF1alpha) in specific tissues, based on the differing expression profile of c1ORF12 and EGLN3 in those tissues.

The Oxford BioMedica clone p1D14 represents Clorf12. The protein sequence encoded by Clorf12 is represented in the public databases by the accession NP_071334 and is described in this patent by Seq ID 89. The nucleotide sequence is represented in the public sequence databases by the accession NM_022051 and is described in this patent by Seq ID 90. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. We find that Clorf12 and a related

human gene EGLN3 (seq ID 85/86) both of which are predicted to be proline hydroxylases, are expressed at differing absolute expression levels in different tissues. For instance, in the hypoxic hepatocyte (6hr) the normalised expression values of EGLN and c1orf12 are 0.015 and 0.0074 respectively, i.e. EGLN being the dominant gene. In contrast, in the neuroblastoma cell line SH-SY5Y, the normalised expression
5 values of EGLN and c1orf12 after 6hr hypoxia are 0.0012 and 0.108 respectively, i.e. c1orf12 being the dominant gene by a large margin. This data demonstrates that c1ORF12 and EGLN3 are not constitutively expressed at an equal amount in different tissues indicating specificity of function. Therefore therapeutic products may be developed based on this data, with the goal of modulating proline hydroxylation of target proteins (such as HIF1alpha) in specific tissues, based on the differing expression
10 profile of c1ORF12 and EGLN3 in those tissues.

The Oxford BioMedica clone p2B1 represents PRAME. The protein sequence encoded by PRAME is represented in the public databases by the accession NP_006106 and is described in this patent by Seq ID 87. The nucleotide sequence is represented in the public sequence databases by the accession NM_006115 and is described in this patent by Seq ID 88. Hypoxia is an important feature of several
15 diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, PRAME is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient. PRAME is a well-known tumour-associated antigen. Our surprising demonstration of
20 its hypoxia-regulation provides for an important diagnostic test to distinguish false-positive results. In addition, we show the relevance of PRAME to hypoxia-related functions of tumours such as angiogenesis.

The Oxford BioMedica clones p1D17 and p1P14 represent Semaphorin 4b. The protein sequence encoded by Semaphorin 4b is represented in the public databases by the accession BAB21836 and is
25 described in this patent by Seq ID 91. The nucleotide sequence is represented in the public sequence databases by the accession AB051532 and is described in this patent by Seq ID 92. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. The response of renal epithelial cells to hypoxia is pertinent to kidney failure, especially regarding the
30 medullary tissue. Semaphorin 4b is preferentially induced by hypoxia in renal epithelial cells. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Semaphorin 4b is induced in macrophages activated by LPS and gamma interferon.. Semaphorin 4b is also induced by

the the presence of reactive oxygen species. We expect it to have a pro-inflammatory role, and its inhibition may have an anti-inflammatory effect. We have cited elsewhere in this specification that a plexin is hypoxia-regulated, and we propose a functional relationship between these two molecules. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of
5 5 patients with either ovarian or breast cancer, Semaphorin 4b is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient. Semaphorin 4b is also induced in response to superoxide radicals, as found in various disease states, implying utility. Semaphorin 4b is predicted to function in modulating several cellular processes key to human disease, including angiogenesis, inflammation, immune cell migration and tissue remodelling. Other Semaphorins including Semaphorin
10 E, which are induced in response to hypoxia will also be implicated in these disease processes and have utility as described for Semaphorin 4b.

The Oxford BioMedica clone p1C24 represents SLC25A19. The protein sequence encoded by SLC25A19 is represented in the public databases by the accession NP_068380 and is described in this patent by Seq ID 93. The nucleotide sequence is represented in the public sequence databases by the accession
15 NM_021734 and is described in this patent by Seq ID 94. SLC25A19 transports deoxynucleotides into mitochondria and is therefore essential for mtDNA synthesis. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1D3 represents Serine carboxypeptidase 1. The protein sequence encoded
20 by Serine carboxypeptidase 1 is represented in the public databases by the accession NP_067639 and is described in this patent by Seq ID 95. The nucleotide sequence is represented in the public sequence databases by the accession NM_021626 and is described in this patent by Seq ID 96. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.
25 Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Serine carboxypeptidase 1 is repressed in macrophages activated by LPS and gamma interferon. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the
30 pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Serine carboxypeptidase 1 is induced in macrophages activated by TNFalpha. Increased serine carboxypeptidase activity in glial cells has been shown to result in neurological abnormalities, due to the degradation of essential neuro-active factors.

Similarly, peripheral neurological disease could result from such activity in macrophages. Our demonstration of hypoxia regulation of serine carboxypeptidase activity opens a route for diagnosis and treatment of these diseases. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Serine carboxypeptidase 1 is
5 down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1E14 represents an unknown mRNA (schizophrenia-linked). The protein sequence encoded by the unknown mRNA (schizophrenia-linked) is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AY010112 and is described in this patent by Seq ID 98. Hypoxia is an important feature
10 of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. TNFalpha is an inflammatory
15 cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Unknown mRNA (schizophrenia-linked) is induced in macrophages activated by TNFalpha. There are many enzymic activities that can give rise to neurological abnormalities, and
20 their hypoxia regulation is pertinent to the diagnosis and treatment of such diseases, including schizophrenia.

The Oxford BioMedica clone p1E20 represents Myo-inositol monophosphatase A3. The protein sequence encoded by Myo-inositol monophosphatase A3 is represented in the public databases by the accession AAK52336 and is described in this patent by Seq ID 99. The nucleotide sequence is represented in the
25 public sequence databases by the accession NM_017813 and is described in this patent by Seq ID 100. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. As referred to elsewhere in this specification, we have found several components of the phosphatidylinositol second messenger system to be hypoxia-regulated. This system has profound effects
30 which are relevant to many diseases with known associations with hypoxia and ischaemia. Local and transient ischaemia is relevant to such diseases as rheumatoid arthritis and atherosclerosis, and also potentially to such diseases as schizophrenia and bi-polar disorder. It is instructive that lithium, which is a well-recognised treatment for affective disorders, appears to operate via the phosphatidylinositol system [Pettegrew et al 2001, Bipolar Disord 3:189-201]. Macrophages are key to several diseases involving

hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Myo-inositol monophosphatase A3 is repressed in macrophages activated by LPS and gamma interferon.

- 5 The Oxford BioMedica clone p2A24 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA521314 and is described in this patent by Seq ID 102. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic,
10 prognostic and diagnostic products.

The Oxford BioMedica clone p1E17 represents Hypothetical protein FLJ31668. The protein sequence encoded by Hypothetical protein FLJ31668 is represented in the public databases by the accession BAB71124 and is described in this patent by Seq ID 103. The nucleotide sequence is represented in the public sequence databases by the accession AK056230 and is described in this patent by Seq ID 104.

- 15 Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

- The Oxford BioMedica clone p1E19 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is
20 represented in the public sequence databases by the accession R51835 and is described in this patent by Seq ID 106. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, the EST represented by Seq ID 106 is
25 up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

- The Oxford BioMedica clone p1E15 represents cDNA YI27F12. The protein sequence encoded by cDNA YI27F12 is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AF075018 and is described in this patent by Seq ID 108. Hypoxia is an important feature of several diseases, and genes that respond to this
30 stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. The cDNA YI27F12 is induced in macrophages treated with the inhibitory cytokine IL-10. The cDNA YI27F12 is repressed in macrophages activated by IL-17. We expect the product of cDNA YI27F12 to have an anti-inflammatory role.

The Oxford BioMedica clone p1E11 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession R69248 and is described in this patent by Seq ID 110. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus
5 are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1E23 represents cDNA FLJ14041 fis, clone HEMBA1005780. The protein sequence encoded by cDNA FLJ14041 fis, clone HEMBA1005780 is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence
10 databases by the accession AK024103 and is described in this patent by Seq ID 112. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1E21 represents Glutamate-cysteine ligase, modifier subunit. The protein sequence encoded by Glutamate-cysteine ligase, modifier subunit is represented in the public databases
15 by the accession NP_002052 and is described in this patent by Seq ID 113. The nucleotide sequence is represented in the public sequence databases by the accession NM_002061 and is described in this patent by Seq ID 114. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Glutamate-cysteine ligase is the rate-limiting enzyme of glutathione
20 synthesis, and this enzyme is relevant to cell survival under stress. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Glutamate-cysteine ligase, modifier subunit is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1D23 represents PTEN. The protein sequence encoded by PTEN is
25 represented in the public databases by the accession NP_000305 and is described in this patent by Seq ID 115. The nucleotide sequence is represented in the public sequence databases by the accession NM_000314 and is described in this patent by Seq ID 116. PTEN is a member of the mixed function, serine/threonine/tyrosine phosphatase subfamily of protein phosphatases. Its physiological substrates, however, are primarily 3-phosphorylated inositol phospholipids, which are products of phosphoinositide
30 3-kinases [Downes et al 2001, Biochem Soc Trans 29:846-51]. Hypoxia-regulation of this gene is a further element in the hypoxic regulation of this important second messenger system. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1D24 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession T73780 and is described in this patent by Seq ID 118. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus
5 are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. The response of renal epithelial cells to hypoxia is pertinent to kidney failure, especially regarding the medullary tissue. The EST represented by Seq ID 118 is preferentially induced by hypoxia in renal epithelial cells. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which
10 have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. The EST represented by Seq ID 118 is induced in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clones p1D22 and p1G5 represent MAX-interacting protein 1. The protein sequence encoded by MAX-interacting protein 1 is represented in the public databases by the accession
15 NP_005953 and is described in this patent by Seq ID 119 and 279. The nucleotide sequence is represented in the public sequence databases by the accession NM_005962 and is described in this patent by Seq ID 120 and 280. MAX-interacting protein 1 is a negative regulator of myc oncoprotein with tumor suppressor properties. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic,
20 prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. MAX-interacting protein 1 is repressed in macrophages activated by LPS and gamma interferon.

25 The Oxford BioMedica clone p1E2 represents Mannosidase, alpha, class 1A, member 1. The protein sequence encoded by Mannosidase, alpha, class 1A, member 1 is represented in the public databases by the accession NP_005898 and is described in this patent by Seq ID 121. The nucleotide sequence is represented in the public sequence databases by the accession NM_005907 and is described in this patent by Seq ID 122. Hypoxia is an important feature of several diseases, and genes that respond to this
30 stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Mannosidase, alpha, class 1A, member 1 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1E1 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA446361 and is described in this patent by Seq ID 124. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. The EST represented by Seq ID 124 is repressed in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1E4 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA931411 and is described in this patent by Seq ID 126. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. The EST represented by Seq ID 126 is repressed in macrophages activated by LPS and gamma interferon. We expect this gene product to have an anti-inflammatory role. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, the EST represented by Seq ID 126 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1D18 represents cDNA FLJ13443 fis, clone PLACE1002853. The protein sequence encoded by cDNA FLJ13443 fis, clone PLACE1002853 is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AK023505 and is described in this patent by Seq ID 128. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. The cDNA FLJ13443 fis, clone PLACE1002853 is repressed in macrophages activated by LPS and gamma interferon. We expect it to have an anti-inflammatory role.

The Oxford BioMedica clone p1D21 represents Hypothetical protein FLJ22622. The protein sequence encoded by Hypothetical protein FLJ22622 is represented in the public databases by the accession BAB15424 and is described in this patent by Seq ID 129. The nucleotide sequence is represented in the public sequence databases by the accession NM_025151 and is described in this patent by Seq ID 130.

5 Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. The response of renal epithelial cells to hypoxia is pertinent to kidney failure, especially regarding the medullary tissue. Hypothetical protein FLJ22622 is preferentially induced by hypoxia in renal epithelial cells. Macrophages are key to several diseases involving hypoxia, and contribute to
10 inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Hypothetical protein FLJ22622 is repressed in macrophages activated by LPS and gamma interferon. We expect it to have an anti-inflammatory role. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast
15 cancer, Hypothetical protein FLJ22622 is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1C22 represents CD84-H1. The protein sequence encoded by CD84-H1 is represented in the public databases by the accession AAK69052 and is described in this patent by Seq ID 131. The nucleotide sequence is represented in the public sequence databases by the accession AF275725
20 and is described in this patent by Seq ID 132. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1C23 represents Hypothetical protein FLJ12832. The protein sequence encoded by Hypothetical protein FLJ12832 is represented in the public databases by the accession
25 XP_043394 and is described in this patent by Seq ID 133. The nucleotide sequence is represented in the public sequence databases by the accession AK022894 and is described in this patent by Seq ID 134. Hypothetical protein FLJ12832 is a putative ubiquitin as it shows high structural similarity to ubiquitin C and contains a ubiquitin domain. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of
30 therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1D11 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA251748 and is described in this patent by Seq ID 136. Hypoxia is an important feature of several diseases, and genes that respond to this

stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clones p1E3 and p1F16 represent CYP1B1. The protein sequence encoded by CYP1B1 is represented in the public databases by the accession NP_000095 and is described in this patent by Seq ID 137 and 325. The nucleotide sequence is represented in the public sequence databases by the accession NM_000104 and is described in this patent by Seq ID 138 and 326. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD. CYP1B1 is preferentially induced by hypoxia in monocytes or macrophages and a restricted number of other cell types. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. CYP1B1 is repressed in macrophages activated by LPS and gamma interferon. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. CYP1B1 is induced in macrophages activated by TNFalpha. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, CYP1B1 is up-regulated and also down regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1D20 represents Hypothetical protein KIAA1125. The protein sequence encoded by Hypothetical protein KIAA1125 is represented in the public databases by the accession XP_012932 and is described in this patent by Seq ID 139. The nucleotide sequence is represented in the public sequence databases by the accession AB032951 and is described in this patent by Seq ID 140. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1E5 represents Hepcidin antimicrobial peptide. The protein sequence encoded by Hepcidin antimicrobial peptide is represented in the public databases by the accession

NP_066998 and is described in this patent by Seq ID 141. The nucleotide sequence is represented in the public sequence databases by the accession NM_021175 and is described in this patent by Seq ID 142. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hepatocytes are the main cell type of the liver and genes that are induced in response to hypoxia in this cell type are relevant to development of diagnostics and therapeutics towards liver diseases involving hypoxia, including cirrhosis. Hepcidin antimicrobial peptide is preferentially induced by hypoxia in hepatocytes. Hepcidin antimicrobial peptide is induced in macrophages treated with the inhibitory cytokine IL-10. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Hepcidin antimicrobial peptide is repressed in macrophages activated by TNFalpha.

The Oxford BioMedica clone p1D19 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession R68736 and is described in this patent by Seq ID 144. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. The EST represented by Seq ID 144 is induced in macrophages activated by LPS and gamma interferon and is also induced in macrophages activated by IL-15. We expect the gene product relevant to the EST represented by Seq ID 144 to have a pro-inflammatory role, and its inhibition may have an anti-inflammatory effect.

The Oxford BioMedica clone p2A15 represents Sialyltransferase. The protein sequence encoded by Sialyltransferase is represented in the public databases by the accession NP_006447 and is described in this patent by Seq ID 145. The nucleotide sequence is represented in the public sequence databases by the accession NM_006456 and is described in this patent by Seq ID 146. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1114 represents cDNA DKFZp564D016. The protein sequence encoded by cDNA DKFZp564D016 is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AL050021 and is described in

this patent by Seq ID 148. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p112 represents cDNA FLJ11302 fis, clone PLACE1009971. The protein
5 sequence encoded by cDNA FLJ11302 fis, clone PLACE1009971 is not represented in the public
databases by a protein accession. The nucleotide sequence is represented in the public sequence databases
by the accession AK002164 and is described in this patent by Seq ID 150. Hypoxia is an important
feature of several diseases, and genes that respond to this stimulus are therefore implicated in the
pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.
10 Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In
these, macrophages are frequently activated by cytokines, which have been shown to be present at disease
sites, so gene expression responses to both hypoxia and cytokines are especially relevant. The cDNA
FLJ11302 fis, clone PLACE1009971 is repressed in macrophages activated by LPS and gamma
interferon. We expect it to have an anti-inflammatory role.

15 The Oxford BioMedica clone p112 represents Hypothetical protein MGC4549. The protein sequence
encoded by Hypothetical protein MGC4549 is represented in the public databases by the accession
XP_032794 and is described in this patent by Seq ID 151. The nucleotide sequence is represented in the
public sequence databases by the accession NM_032377 and is described in this patent by Seq ID 152.
Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore
20 implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic
products. Hypothetical protein MGC4549 is induced in macrophages treated with the inhibitory cytokine
IL-10. Hypothetical protein MGC4549 is repressed in macrophages activated by IL-17 and is also
repressed in macrophages activated by IL-15. We expect it to have an anti-inflammatory role.

The Oxford BioMedica clone p113 represents ELMO2. The protein sequence encoded by ELMO2 is
25 represented in the public databases by the accession AAL14467 and is described in this patent by Seq ID
153. The nucleotide sequence is represented in the public sequence databases by the accession
XM_012933 and is described in this patent by Seq ID 154. Hypoxia is an important feature of several
diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have
utility in the design of therapeutic, prognostic and diagnostic products. This gene has been shown recently
30 to promote phagocytosis and cell shape changes [Gumienny et al 2001, Cell 107:27-41]. These functions
are typical of the macrophage, and are likely to play a role in macrophage-associated diseases.

The Oxford BioMedica clone p110 represents an unannotated EST. The protein sequence encoded by
this EST is not represented in the public databases by a protein accession. The nucleotide sequence is

represented in the public sequence databases by the accession AA420992 and is described in this patent by Seq ID 156. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

5 The Oxford BioMedica clone p1H18 represents Ubiquitin specific protease 7. The protein sequence encoded by Ubiquitin specific protease 7 is represented in the public databases by the accession NP_003461 and is described in this patent by Seq ID 157. The nucleotide sequence is represented in the public sequence databases by the accession NM_003470 and is described in this patent by Seq ID 158. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore
10 implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Ubiquitin specific protease 7 is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient. We expect decreased activity of the gene product to have an anti-tumour effect.

15 The Oxford BioMedica clone p1H24 represents Nucleolar phosphoprotein Nopp34. The protein sequence encoded by Nucleolar phosphoprotein Nopp34 is represented in the public databases by the accession NP_115766 and is described in this patent by Seq ID 159. The nucleotide sequence is represented in the public sequence databases by the accession NM_032390 and is described in this patent by Seq ID 160. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore
20 implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1E22 represents cDNA FLJ13618 fis, clone PLACE1010925. The protein sequence encoded by cDNA FLJ13618 fis, clone PLACE1010925 is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases
25 by the accession AK023680 and is described in this patent by Seq ID 162. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease
30 sites, so gene expression responses to both hypoxia and cytokines are especially relevant. The cDNA FLJ13618 fis, clone PLACE1010925 is induced in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1H21 represents Hypothetical protein FLJ13511. The protein sequence encoded by Hypothetical protein FLJ13511 is represented in the public databases by the accession

NP_149014 and is described in this patent by Seq ID 163. The nucleotide sequence is represented in the public sequence databases by the accession NM_033025 and is described in this patent by Seq ID 164. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD. Hypothetical protein FLJ13511 is preferentially induced by hypoxia in monocytes or macrophages and a restricted number of other cell types. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Hypothetical protein FLJ13511 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1I1 represents Ribosomal RNA intergenic spacer. The protein sequence encoded by Ribosomal RNA intergenic spacer is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA664228 and is described in this patent by Seq ID 166. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1H14 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession R44397 and is described in this patent by Seq ID 168. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1H11 represents Carboxypeptidase M. The protein sequence encoded by Carboxypeptidase M is represented in the public databases by the accession NP_001865 and is described in this patent by Seq ID 169. The nucleotide sequence is represented in the public sequence databases by the accession NM_001874 and is described in this patent by Seq ID 170. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1H17 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is

represented in the public sequence databases by the accession W87747 and is described in this patent by Seq ID 172. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, the EST represented by Seq ID 172 is up-regulated and also down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1H12 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA973568 and is described in this patent by Seq ID 174. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1H7 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession T98529 and is described in this patent by Seq ID 176. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1H15 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA022679 and is described in this patent by Seq ID 178. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, the EST represented by Seq ID 178 is up-regulated and also down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1H20 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession H17921 and is described in this patent by Seq ID 180. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and

diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, the EST represented by Seq ID 180 is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient. We expect decreased activity of the gene product to have an anti-tumour effect.

- 5 The Oxford BioMedica clone p1H8 represents ABL. The protein sequence encoded by ABL is represented in the public databases by the accession NP_009297 and is described in this patent by Seq ID 181. The nucleotide sequence is represented in the public sequence databases by the accession NM_007313 and is described in this patent by Seq ID 182. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have
- 10 utility in the design of therapeutic, prognostic and diagnostic products. ABL is induced in macrophages treated with the inhibitory cytokine IL-10. ABL is repressed in macrophages activated by IL-17 and is also repressed in macrophages activated by IL-15. We expect it to have an anti-inflammatory role. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, ABL is up-regulated in the malignant tissue as compared
- 15 to adjacent normal tissue in at least one patient.

- The Oxford BioMedica clone p1H16 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession W91958 and is described in this patent by Seq ID 184. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus
- 20 are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, the EST represented by Seq ID 184 is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

- The Oxford BioMedica clone p1H9 represents an unannotated EST. The protein sequence encoded by this
- 25 EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession R63694 and is described in this patent by Seq ID 186. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

- 30 The Oxford BioMedica clone p1H23 represents Hypothetical protein FLJ21094. The protein sequence encoded by Hypothetical protein FLJ21094 is represented in the public databases by the accession AAH14003 and is described in this patent by Seq ID 187. The nucleotide sequence is represented in the public sequence databases by the accession AK024747 and is described in this patent by Seq ID 188.

Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1H10 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA909912 and is described in this patent by Seq ID 190. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

10 The Oxford BioMedica clone p1H6 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession T99032 and is described in this patent by Seq ID 192. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. The EST represented by Seq ID 192 is induced in macrophages treated with the inhibitory cytokine IL-10. The EST represented by Seq ID 192 is repressed in macrophages activated by IL-15. We expect it to have an anti-inflammatory role.

The Oxford BioMedica clone p1H13 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession H52503 and is described in this patent by Seq ID 194. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites; so gene expression responses to both hypoxia and cytokines are especially relevant. The EST represented by Seq ID 194 is repressed in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1H19 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA127017 and is described in this patent by Seq ID 196. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage

infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, the EST represented by the Seq ID 196 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1G22 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession R38647 and is described in this patent by Seq ID 198. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Endothelial cells are key to angiogenesis, a process implicated in several diseases associated with hypoxia, including cancer and rheumatoid arthritis. The EST represented by Seq ID 198 is preferentially induced by hypoxia in endothelial cells. We expect this gene product to have a pro-angiogenic effect, and its inhibition to have an anti-angiogenic effect.

The Oxford BioMedica clone p1G21 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession T87233 and is described in this patent by Seq ID 200. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1H1 represents Hypothetical protein FLJ10826. The protein sequence encoded by Hypothetical protein FLJ10826 is represented in the public databases by the accession BAB14226 and is described in this patent by Seq ID 201. The nucleotide sequence is represented in the public sequence databases by the accession NM_018233 and is described in this patent by Seq ID 202. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1G20 represents cDNA YO23H03. The protein sequence encoded by cDNA YO23H03 is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AF075053 and is described in this patent by Seq ID 204. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responds to both hypoxia and

cytokines are especially relevant. The cDNA YO23H03 is repressed in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1H5 represents Hypothetical protein FLJ22690. The protein sequence encoded by Hypothetical protein FLJ22690 is represented in the public databases by the accession
5 NP_078987 and is described in this patent by Seq ID 205. The nucleotide sequence is represented in the public sequence databases by the accession NM_024711 and is described in this patent by Seq ID 206. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Endothelial cells are key to angiogenesis, a process implicated in several diseases associated
10 with hypoxia, including cancer and rheumatoid arthritis. Hypothetical protein FLJ22690 is preferentially induced by hypoxia in endothelial cells. We expect this gene product to have a pro-angiogenic effect, and its inhibition to have an anti-angiogenic effect. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both
15 hypoxia and cytokines are especially relevant. Hypothetical protein FLJ22690 is induced in macrophages activated by IL-15.

The Oxford BioMedica clone p1G19 represents Mitochondrion sequence. The protein sequence encoded by Mitochondrion sequence is represented in the public databases by the accession AAH05845 and is described in this patent by Seq ID 207. The nucleotide sequence is represented in the public sequence
20 databases by the accession BC005845 and is described in this patent by Seq ID 208. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, the Mitochondrion sequence represented by Seq ID 208 is
25 down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1H2 represents Fatty acid binding protein 5. The protein sequence encoded by Fatty acid binding protein 5 is represented in the public databases by the accession NP_001435 and is described in this patent by Seq ID 209. The nucleotide sequence is represented in the public sequence databases by the accession NM_001444 and is described in this patent by Seq ID 210.
30 Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell

types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD. Fatty acid binding protein 5 is preferentially induced by hypoxia in monocytes or macrophages. Crucially and very recently, Fatty acid binding protein 5 expressed in macrophages has been shown to play a very important role in the development of atherosclerotic plaques [Layne et al 2001, FASEB J 15:2733-5]. Our demonstration of hypoxic-regulation of this gene not only makes clear how this gene can participate in disease initiation and progression, but provides for a potential route to diagnosis and therapy of atherosclerosis. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Fatty acid binding protein 5 is repressed in macrophages activated by TNFalpha.

The Oxford BioMedica clone p1G18 represents Mitochondrion sequence. The protein sequence encoded by Mitochondrion sequence is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession BC001612 and is described in this patent by Seq ID 212. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. The Mitochondrion sequence represented by Seq ID 212 is repressed in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1H4 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA679939 and is described in this patent by Seq ID 214. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. The EST represented by Seq ID 214 is repressed in macrophages

activated by IL-17. We expect it to have an anti-inflammatory role. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, the EST represented by the Seq ID 214 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

- 5 The Oxford BioMedica clone p1H3 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA630167 and is described in this patent by Seq ID 216. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, the EST represented by Seq ID 216 is up-regulated and also down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

- The protein sequence encoded by BCL2/adenovirus E1B 19kD-interacting protein 3-like is represented in the public databases by the accession NP_004322 and is described in this patent by Seq ID 217. The nucleotide sequence is represented in the public sequence databases by the accession NM_004331 and is described in this patent by Seq ID 218. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

- 20 The protein sequence encoded by SLC2A1 is represented in the public databases by the accession NP_006507 and is described in this patent by Seq ID 219. The nucleotide sequence is represented in the public sequence databases by the accession NM_006516 and is described in this patent by Seq ID 220. SLC2A1 is a glucose transporter gene and is also known as GLUT1. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

- The Oxford BioMedica clone p1P3 represents PDGFB. The protein sequence encoded by PDGFB is represented in the public databases by the accession NP_148937 and is described in this patent by Seq ID 221. The nucleotide sequence is represented in the public sequence databases by the accession NM_033016 and is described in this patent by Seq ID 222. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene

expression responses to both hypoxia and cytokines are especially relevant. PDGFB is induced in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clones p1A8 and p1A9 represent Lactate dehydrogenase A. The protein sequence encoded by Lactate dehydrogenase A is represented in the public databases by the accession NP_005557
5 and is described in this patent by Seq ID 223. The nucleotide sequence is represented in the public sequence databases by the accession NM_005566 and is described in this patent by Seq ID 224. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In
10 these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Lactate dehydrogenase A is repressed in macrophages activated by LPS and gamma interferon and is also repressed in macrophages activated by IL-15. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases
15 including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Lactate dehydrogenase A is induced in macrophages activated by TNFalpha. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Lactate dehydrogenase A is up-regulated in the malignant tissue as compared to
20 adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1B17 represents Tissue factor. The protein sequence encoded by Tissue factor is represented in the public databases by the accession NP_001984 and is described in this patent by Seq ID 225. The nucleotide sequence is represented in the public sequence databases by the accession NM_001993 and is described in this patent by Seq ID 226. Hypoxia is an important feature of several
25 diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Tissue factor is preferentially induced by hypoxia in mammary epithelial cells. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both
30 hypoxia and cytokines are especially relevant. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Tissue factor is induced in macrophages activated by TNFalpha. Tissue factor is the primary initiator of

blood coagulation with structural homology to the cytokine receptor family, and has been implicated in various vascular processes including metastasis, angiogenesis, and atherosclerosis. Our demonstration of hypoxic regulation leads to a clear understanding of the possibility of intervention in disease by modulation of Tissue factor activity.

- 5 The Oxford BioMedica clone p1O20 represents VEGF. The protein sequence encoded by VEGF is represented in the public databases by the accession NP_003367 and is described in this patent by Seq ID 227. The nucleotide sequence is represented in the public sequence databases by the accession NM_003376 and is described in this patent by Seq ID 228. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have
10 utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, VEGF is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

- The Oxford BioMedica clone p1B2 represents N-myc downstream regulated. The protein sequence
15 encoded by N-myc downstream regulated is represented in the public databases by the accession NP_006087 and is described in this patent by Seq ID 229. The nucleotide sequence is represented in the public sequence databases by the accession NM_006096 and is described in this patent by Seq ID 230. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic
20 products. N-myc downstream regulated is preferentially induced by hypoxia in mammary epithelial cells.

- The Oxford BioMedica clone p1B3 represents Proline 4-hydroxylase, alpha polypeptide I. The protein sequence encoded by Proline 4-hydroxylase, alpha polypeptide I is represented in the public databases by the accession NP_000908 and is described in this patent by Seq ID 231. The nucleotide sequence is represented in the public sequence databases by the accession NM_000917 and is described in this patent
25 by Seq ID 232. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responds to both hypoxia and
30 cytokines are especially relevant. Proline 4-hydroxylase, alpha polypeptide I is repressed in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Proline

4-hydroxylase, alpha polypeptide I is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The protein sequence encoded by BCL2/adenovirus E1B-interacting protein 3 is represented in the public databases by the accession NP_004043 and is described in this patent by Seq ID 233. The nucleotide
5 sequence is represented in the public sequence databases by the accession NM_004052 and is described in this patent by Seq ID 234. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clones p1B18 and p1B19 represent Plasminogen activator inhibitor, type 1. The
10 protein sequence encoded by Plasminogen activator inhibitor, type 1 is represented in the public databases by the accession NP_000593 and is described in this patent by Seq ID 235. The nucleotide sequence is represented in the public sequence databases by the accession NM_000602 and is described in this patent by Seq ID 236. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic,
15 prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Plasminogen activator inhibitor, type 1 is induced in macrophages activated by LPS and gamma interferon. Plasminogen activator inhibitor, type 1 is repressed in
20 macrophages activated by IL-17. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Plasminogen activator inhibitor, type 1 is induced in macrophages activated by TNFalpha. Hypoxia is frequently found in human
25 tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Plasminogen activator inhibitor, type 1 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1N17 represents COX-2. The protein sequence encoded by COX-2 is represented in the public databases by the accession NP_000954 and is described in this patent by Seq ID
30 237. The nucleotide sequence is represented in the public sequence databases by the accession NM_000963 and is described in this patent by Seq ID 238. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. COX-2 is preferentially induced by hypoxia in mammary epithelial cells. Macrophages are key to several diseases involving hypoxia, and

contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. COX-2 is induced in macrophages activated by LPS and gamma interferon. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. COX-2 is induced in macrophages activated by TNFalpha. In view of the known role of COX-2 in prostaglandin synthesis and tumour progression, its induction by hypoxia has profound clinical implications, and clear utility in diagnosis and therapy.

10 Hypoxia is frequently found in human tumours where macrophage infiltrates are also found.

The Oxford BioMedica clone p1A24 represents Metallothionein 1H. The protein sequence encoded by Metallothionein 1H is represented in the public databases by the accession NP_005942 and is described in this patent by Seq ID 239. The nucleotide sequence is represented in the public sequence databases by the accession NM_005951 and is described in this patent by Seq ID 240. Metallothioneins can act as an antioxidant and free-radical scavenger and are therefore protective against cell death in hypoxia. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hepatocytes are the main cell type of the liver and genes which are induced in response to hypoxia in this cell type are relevant to development of diagnostics and therapeutics towards liver diseases involving hypoxia, including cirrhosis. Metallothionein 1H is preferentially induced by hypoxia in hepatocytes. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Metallothionein 1H is induced in macrophages activated by LPS and gamma interferon.

15

20

25 The protein sequence encoded by Metallothionein 1L is represented in the public databases by the accession NP_002441 and is described in this patent by Seq ID 241. The nucleotide sequence is represented in the public sequence databases by the accession NM_002450 and is described in this patent by Seq ID 242. Metallothioneins can act as an antioxidant and free-radical scavenger and are therefore protective against cell death in hypoxia. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

30

The Oxford BioMedica clone p1B1 represents Metallothionein 1G. The protein sequence encoded by Metallothionein 1G is represented in the public databases by the accession NP_005941 and is described in this patent by Seq ID 243. The nucleotide sequence is represented in the public sequence databases by the

- accession NM_005950 and is described in this patent by Seq ID 244. Metallothioneins can act as an antioxidant and free-radical scavenger and are therefore protective against cell death in hypoxia. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.
- 5 HIF1alpha and EPAS1 are transcription factors which mediate the response to hypoxia of several genes, and have themselves been implicated in specific diseases. By adenoviral over-expression of HIF1alpha we show augmentation of the hypoxic induction of certain genes, further confirming their status as responsive to hypoxia. Metallothionein 1G has been shown to be induced by hypoxia to a greater degree following adenoviral over-expression of HIF1alpha. Hepatocytes are the main cell type of the liver and
- 10 genes which are induced in response to hypoxia in this cell type are relevant to development of diagnostics and therapeutics towards liver diseases involving hypoxia, including cirrhosis. Metallothionein 1G is preferentially induced by hypoxia in hepatocytes. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene
- 15 expression responses to both hypoxia and cytokines are especially relevant. Metallothionein 1G is induced in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Metallothionein 1G is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.
- 20 The protein sequence encoded by Metallothionein 1E (functional) is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA872383 and is described in this patent by Seq ID 246. Metallothioneins can act as an antioxidant and free-radical scavenger and are therefore protective against cell death in hypoxia. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated
- 25 in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.
- The Oxford BioMedica clones p1A1, p1A2, p1A3 and p1A4 represent SLC2A3. The protein sequence encoded by SLC2A3 is represented in the public databases by the accession NP_008862 and is described in this patent by Seq ID 247. The nucleotide sequence is represented in the public sequence databases by the accession NM_006931 and is described in this patent by Seq ID 248. Hypoxia is an important feature
- 30 of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. SLC2A3 is induced in macrophages treated with the inhibitory cytokine IL-10. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have

utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. SLC2A3 is induced in macrophages activated by TNFalpha.

The Oxford BioMedica clones p1A15, p1A16, p1A17 and p1A18 represent Hexokinase-2. The protein sequence encoded by Hexokinase-2 is represented in the public databases by the accession NP_000180 and is described in this patent by Seq ID 249. The nucleotide sequence is represented in the public sequence databases by the accession NM_000189 and is described in this patent by Seq ID 250. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Hexokinase-2 is repressed in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clones p1B14, p1B15 and p1B16 represent Interleukin 8. The protein sequence encoded by Interleukin 8 is represented in the public databases by the accession NP_000575 and is described in this patent by Seq ID 251. The nucleotide sequence is represented in the public sequence databases by the accession NM_000584 and is described in this patent by Seq ID 252. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Interleukin 8 is induced in macrophages activated by LPS and gamma interferon and is also induced in macrophages activated by IL-17. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Interleukin 8 is induced in macrophages activated by TNFalpha. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Interleukin 8 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clones p1A11 and p1A12 represent GAPDH. The protein sequence encoded by GAPDH is represented in the public databases by the accession NP_002037 and is described in this patent by Seq ID 253. The nucleotide sequence is represented in the public sequence databases by the accession NM_002046 and is described in this patent by Seq ID 254. Hypoxia is an important feature of several

diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. GAPDH is repressed in macrophages activated by LPS and gamma interferon and is also induced in macrophages activated by IL-17 or IL-15. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. GAPDH is induced in macrophages activated by TNFalpha. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, GAPDH is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1A13 represents Phosphoglycerate kinase 1. The protein sequence encoded by Phosphoglycerate kinase 1 is represented in the public databases by the accession NP_000282 and is described in this patent by Seq ID 255. The nucleotide sequence is represented in the public sequence databases by the accession NM_000291 and is described in this patent by Seq ID 256. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Phosphoglycerate kinase 1 is repressed in macrophages activated by LPS and gamma interferon. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Phosphoglycerate kinase 1 is induced in macrophages activated by TNFalpha.

The Oxford BioMedica clone p1A14 represents Enolase 1. The protein sequence encoded by Enolase 1 is represented in the public databases by the accession NP_001419 and is described in this patent by Seq ID 257. The nucleotide sequence is represented in the public sequence databases by the accession NM_001428 and is described in this patent by Seq ID 258. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several

diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Enolase 1 is repressed in macrophages activated by LPS and gamma interferon. TNFalpha is an inflammatory cytokine, which acts
5 on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Enolase 1 is induced in macrophages activated by TNFalpha. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast
10 cancer, Enolase 1 is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1A19 represents Aldolase C. The protein sequence encoded by Aldolase C is represented in the public databases by the accession NP_005156 and is described in this patent by Seq ID 259. The nucleotide sequence is represented in the public sequence databases by the accession
15 NM_005165 and is described in this patent by Seq ID 260. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene
20 expression responses to both hypoxia and cytokines are especially relevant. Aldolase C is induced in macrophages activated by IL-15. Aldolase C is repressed in macrophages activated by IL-15. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and
25 diagnostic products for such inflammatory conditions. Aldolase C is induced in macrophages activated by TNFalpha. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Aldolase is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1A20 represents Triosephosphate isomerase 1. The protein sequence
30 encoded by Triosephosphate isomerase 1 is represented in the public databases by the accession NP_000356 and is described in this patent by Seq ID 261. The nucleotide sequence is represented in the public sequence databases by the accession NM_000365 and is described in this patent by Seq ID 262. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic

- products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Triosephosphate isomerase 1 is repressed in macrophages activated by LPS and gamma
- 5 interferon and is also repressed in macrophages activated by IL-15. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Triosephosphate isomerase 1 is induced in macrophages activated by TNFalpha.
- 10 The Oxford BioMedica clone p1A23 represents Metallothionein 2A. The protein sequence encoded by Metallothionein 2A is represented in the public databases by the accession NP_005944 and is described in this patent by Seq ID 265. The nucleotide sequence is represented in the public sequence databases by the accession NM_005953 and is described in this patent by Seq ID 266. Metallothioneins can act as an antioxidant and free-radical scavenger and are therefore protective against cell death in hypoxia. Hypoxia
- 15 is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. HIF1alpha and EPAS1 are transcription factors which mediate the response to hypoxia of several genes, and have themselves been implicated in specific diseases. By adenoviral over-expression of HIF1alpha we show augmentation of the hypoxic induction of certain genes, further confirming their status as
- 20 responsive to hypoxia. Metallothionein 2A has been shown to be induced by hypoxia to a greater degree following adenoviral over-expression of HIF1alpha. Hepatocytes are the main cell type of the liver and genes which are induced in response to hypoxia in this cell type are relevant to development of diagnostics and therapeutics towards liver diseases involving hypoxia, including cirrhosis. Metallothionein 2A is preferentially induced by hypoxia in hepatocytes. Macrophages are key to several
- 25 diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Metallothionein 2A is induced in macrophages activated by LPS and gamma interferon and also induced in macrophages activated by IL-15. Hypoxia is frequently found in human tumours where macrophage infiltrates are also
- 30 found. In a series of 5 patients with either ovarian or breast cancer, Metallothionein 2A is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clones p1B20 and p1B21 represent Osteopontin. The protein sequence encoded by Osteopontin is represented in the public databases by the accession NP_000573 and is described in this patent by Seq ID 267. The nucleotide sequence is represented in the public sequence databases by the

accession NM_000582 and is described in this patent by Seq ID 268. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated
5 by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD. Osteopontin is preferentially induced by hypoxia in monocytes or macrophages and a restricted number of other cell types. Osteopontin has been shown recently to play a role in autoimmune disease
10 [Chabas et al, 2001, Science 294: 1731-5]. We present a new association between the hypoxic response and autoimmune disease. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Osteopontin is repressed in macrophages activated by LPS and gamma interferon.
15 Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Osteopontin is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clones p1C17 and p1C18 represent Granulin. The protein sequence encoded by Granulin is represented in the public databases by the accession NP_002078 and is described in this
20 patent by Seq ID 269. The nucleotide sequence is represented in the public sequence databases by the accession NM_002087 and is described in this patent by Seq ID 270. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are
25 frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Granulin is repressed in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Granulin is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one
30 patient. The up-regulation of Granulin, which is a known growth factor, is a clinically significant feature of the hypoxic response with clear diagnostic and therapeutic utility.

The Oxford BioMedica clone p1D8 represents Hypoxia-inducible protein 2. The protein sequence encoded by Hypoxia-inducible protein 2 is represented in the public databases by the accession NP_037464 and is described in this patent by Seq ID 271. The nucleotide sequence is represented in the

public sequence databases by the accession NM_013332 and is described in this patent by Seq ID 272. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia-inducible protein 2 is induced in macrophages treated with the inhibitory cytokine IL-
5 10. Hypoxia-inducible protein 2 is repressed in macrophages activated by IL-17 and is also repressed in macrophages activated by IL-15.

The Oxford BioMedica clone p1A10 represents Enolase 2. The protein sequence encoded by Enolase 2 is represented in the public databases by the accession NP_001966 and is described in this patent by Seq ID 273. The nucleotide sequence is represented in the public sequence databases by the accession
10 NM_001975 and is described in this patent by Seq ID 274. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene
15 expression responses to both hypoxia and cytokines are especially relevant. Enolase 2 is repressed in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Enolase 2 is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient. We expect it to have an anti-inflammatory role.

20 The Oxford BioMedica clone p1G24 represents Glycogen synthase 1. The protein sequence encoded by Glycogen synthase 1 is represented in the public databases by the accession NP_002094 and is described in this patent by Seq ID 275. The nucleotide sequence is represented in the public sequence databases by the accession NM_002103 and is described in this patent by Seq ID 276. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis
25 and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Glycogen synthase 1 is repressed in macrophages activated by IL-17 and is also repressed in macrophages activated by IL-15.

30 The Oxford BioMedica clone p1G23 represents ALCAM. The protein sequence encoded by ALCAM is represented in the public databases by the accession NP_001618 and is described in this patent by Seq ID 277. The nucleotide sequence is represented in the public sequence databases by the accession NM_001627 and is described in this patent by Seq ID 278. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have

utility in the design of therapeutic, prognostic and diagnostic products. In view of the recently-discovered role of ALCAM in angiogenesis [Ohneda et al, 2001, Blood 2001 Oct 1;98(7):2134-42], our demonstration of hypoxic regulation of ALCAM has great clinical significance in the treatment and diagnosis of vascular disease and cancer.

- 5 The Oxford BioMedica clone p1G7 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession BC008022 and is described in this patent by Seq ID 282. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. The EST represented by Seq ID 282 is repressed in macrophages activated by LPS and gamma interferon. We expect the product of EST represented by Seq ID 282 to have an anti-inflammatory role.

- The Oxford BioMedica clone p2A23 represents Chitinase 3-like 2. The protein sequence encoded by Chitinase 3-like 2 is represented in the public databases by the accession NP_003991 and is described in this patent by Seq ID 283. The nucleotide sequence is represented in the public sequence databases by the accession NM_004000 and is described in this patent by Seq ID 284. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Chitinase 3-like 2 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

- 25 The Oxford BioMedica clone p1G1 represents BACH1. The protein sequence encoded by BACH1 is represented in the public databases by the accession NP_001177 and is described in this patent by Seq ID 285. The nucleotide sequence is represented in the public sequence databases by the accession NM_001186 and is described in this patent by Seq ID 286. BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. The induction by hypoxia of this known transcriptional repressor and potential oncogene [Cantor et al 2001, Cell 105:149-60] is a very significant finding with profound implications for the diagnosis and treatment of cancer.

The Oxford BioMedica clone p1G15 represents Phosphoglucomutase 1. The protein sequence encoded by Phosphoglucomutase 1 is represented in the public databases by the accession NP_002624 and is described in this patent by Seq ID 287. The nucleotide sequence is represented in the public sequence databases by the accession NM_002633 and is described in this patent by Seq ID 288. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Phosphoglucomutase 1 is induced in macrophages treated with the inhibitory cytokine IL-10.

The Oxford BioMedica clone p1F23 represents Hypothetical protein LOC51014. The protein sequence encoded by Hypothetical protein LOC51014 is represented in the public databases by the accession Q9Y3B3 and is described in this patent by Seq ID 289. The nucleotide sequence is represented in the public sequence databases by the accession AF151867 and is described in this patent by Seq ID 290. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Hypothetical protein LOC51014 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1G8 represents Sin3-associated polypeptide. The protein sequence encoded by Sin3-associated polypeptide is represented in the public databases by the accession NP_003855 and is described in this patent by Seq ID 291. The nucleotide sequence is represented in the public sequence databases by the accession NM_003864 and is described in this patent by Seq ID 292. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1G13 represents ABCA1. The protein sequence encoded by ABCA1 is represented in the public databases by the accession NP_005493 and is described in this patent by Seq ID 293. The nucleotide sequence is represented in the public sequence databases by the accession NM_005502 and is described in this patent by Seq ID 294. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. ABCA1 is repressed in macrophages activated by LPS and gamma interferon. The hypoxia induction of ABCA1, which is known

to be relevant to atherosclerosis [Kielar et al 2001, Clin Chem 47:2089-97], has profound implications for the diagnosis and treatment of this disease.

The Oxford BioMedica clone p1G10 represents SEC24 member A. The protein sequence encoded by SEC24 member A is represented in the public databases by the accession CAA10334 and is described in
5 this patent by Seq ID 295. The nucleotide sequence is represented in the public sequence databases by the accession AJ131244 and is described in this patent by Seq ID 296. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1F24 represents Glia-derived nexin. The protein sequence encoded by
10 Glia-derived nexin is represented in the public databases by the accession AAA35883 and is described in this patent by Seq ID 297. The nucleotide sequence is represented in the public sequence databases by the accession M17783 and is described in this patent by Seq ID 298. Glia-derived nexin is a glycoprotein that functions as a serine protease inhibitor with activity towards thrombin, trypsin and urokinase. It is known to play a role in neuro-degeneration [Seidel et al 1998, Brain Res Mol Brain Res 60:296-300]. Thus the
15 hypoxia induction of this gene is highly significant for the diagnosis and treatment of neuro-degenerative disease. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been
20 shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Glia-derived nexin is induced in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Glia-derived nexin is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

25 The Oxford BioMedica clone p1G2 represents Postsynaptic density-95. The protein sequence encoded by Postsynaptic density-95 is represented in the public databases by the accession NP_001356 and is described in this patent by Seq ID 299. The nucleotide sequence is represented in the public sequence databases by the accession NM_001365 and is described in this patent by Seq ID 300. Postsynaptic density-95 belongs to the MAGUK family of cell junction proteins. Hypoxia is an important feature of
30 several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. The recent demonstration for a possible role for Postsynaptic density-95 in ischaemic pre-conditioning [Tauskela et al 2001,

Neuroscience 107:571-584] underlines the medical significance of our determination of the hypoxic regulation of this gene, and its utility in the diagnosis and treatment of ischaemic disease.

The Oxford BioMedica clone pIG11 represents Tumour protein D52. The protein sequence encoded by Tumour protein D52 is represented in the public databases by the accession NP_005070 and is described
5 in this patent by Seq ID 301. The nucleotide sequence is represented in the public sequence databases by the accession NM_005079 and is described in this patent by Seq ID 302. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. The response of renal epithelial cells to hypoxia is pertinent to kidney failure, especially regarding the medullary tissue. Tumour
10 protein D52 is preferentially induced by hypoxia in renal epithelial cells. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Tumour protein D52 is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient. Our observation of hypoxia-regulation of this tumour-associated protein is highly significant for the diagnosis and treatment of cancer.

15 The Oxford BioMedica clone pIG16 represents p27, Kip1. The protein sequence encoded by p27, Kip1 is represented in the public databases by the accession NP_004055 and is described in this patent by Seq ID 303. The nucleotide sequence is represented in the public sequence databases by the accession NM_004064 and is described in this patent by Seq ID 304. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have
20 utility in the design of therapeutic, prognostic and diagnostic products. The hypoxia regulation of this anti-mitogen has important utility in oncology and angiogenesis [Fouty et al 2001, Am J Respir Cell Mol Biol 25:652-658].

The Oxford BioMedica clone pIG9 represents PI-3-kinase, catalytic, beta polypeptide. The protein sequence encoded by PI-3-kinase, catalytic, beta polypeptide is represented in the public databases by the
25 accession NP_006210 and is described in this patent by Seq ID 305. The nucleotide sequence is represented in the public sequence databases by the accession NM_006219 and is described in this patent by Seq ID 306. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and
30 contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. PI-3-kinase, catalytic, beta polypeptide is repressed in macrophages activated by LPS and gamma interferon. The very recent publication of a role for PI3 kinase in

angiogenesis induced by hypoxic pre-conditioning [Zhu et al 2001, FEBS Lett 508:369-74] re-enforces the clinical utility which we claim for this gene as a result of its hypoxia-induction.

The Oxford BioMedica clone p1G4 represents SLC5A3. The protein sequence encoded by SLC5A3 is represented in the public databases by the accession AAC39548 and is described in this patent by Seq ID 5 307. The nucleotide sequence is represented in the public sequence databases by the accession AF027153 and is described in this patent by Seq ID 308. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. SLC5A3 is over-expressed in the brains of children with Downs Syndrome, and may play a role in brain pathology [Berry et al 1999, J Pediatr 10 135:94-7]. Thus our claims of clinical utility following from hypoxia induction have great medical significance for the diagnosis and treatment of ischaemic and degenerative disease. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. SLC5A3 is repressed in 15 macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1G14 represents Cytohesin binding protein. The protein sequence encoded by Cytohesin binding protein is represented in the public databases by the accession NP_004279 and is described in this patent by Seq ID 309. The nucleotide sequence is represented in the public sequence databases by the accession NM_004288 and is described in this patent by Seq ID 310. Hypoxia is an 20 important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Cytohesin has been shown to modulate PI3-kinase activity [Dierks et al 2001, J Biol Chem 276:37472-81], re-enforcing our claim here and elsewhere in this filing of the relevance to the hypoxic response of pathways controlled by the critical second-messenger PI3.

25 The Oxford BioMedica clones p1A5 and p1A6 represent SLC2A5. The protein sequence encoded by SLC2A5 is represented in the public databases by the accession NP_003030 and is described in this patent by Seq ID 311. The nucleotide sequence is represented in the public sequence databases by the accession NM_003039 and is described in this patent by Seq ID 312. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and 30 have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. SLC2A5 is repressed in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours

where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, SLC2A5 is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

- The Oxford BioMedica clones p1B6, p1B7, p1B8 and p1B9 represent Adipophilin. The protein sequence encoded by Adipophilin is represented in the public databases by the accession NP_001113 and is described in this patent by Seq ID 313. The nucleotide sequence is represented in the public sequence databases by the accession NM_001122 and is described in this patent by Seq ID 314. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD. The hypoxia induction of adipophilin has profound implications for the causation, diagnosis and treatment of atherosclerosis, because this protein plays a key role in the uptake of lipid and foam cell formation [Buechler et al 2001, Biochim Biophys Acta 1532:97-104]. Adipophilin is preferentially induced by hypoxia in monocytes or macrophages and a restricted number of other cell types. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Adipophilin is repressed in macrophages activated by LPS and gamma interferon and is also repressed in macrophages activated by IL-15. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Adipophilin is induced in macrophages activated by TNFalpha. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Adipophilin is up-regulated and also down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.
- The Oxford BioMedica clone p1G17 represents Early development regulator 2. The protein sequence encoded by Early development regulator 2 is represented in the public databases by the accession NP_004418 and is described in this patent by Seq ID 315. The nucleotide sequence is represented in the public sequence databases by the accession NM_004427 and is described in this patent by Seq ID 316. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore

implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Early development regulator 2 is repressed in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1G3 represents B-cell translocation gene 1. The protein sequence encoded by B-cell translocation gene 1 is represented in the public databases by the accession NP_001722 and is described in this patent by Seq ID 317. The nucleotide sequence is represented in the public sequence databases by the accession NM_001731 and is described in this patent by Seq ID 318. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. B-cell translocation gene 1 is induced in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, B-cell translocation gene 1 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1F22 represents Sorting nexin 9. The protein sequence encoded by Sorting nexin 9 is represented in the public databases by the accession NP_057308 and is described in this patent by Seq ID 319. The nucleotide sequence is represented in the public sequence databases by the accession NM_016224 and is described in this patent by Seq ID 320. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Sorting nexin 9 is induced in macrophages activated by TNFalpha.

The Oxford BioMedica clone p1G12 represents Cyclin G2. The protein sequence encoded by Cyclin G2 is represented in the public databases by the accession NP_004345 and is described in this patent by Seq

ID 321. The nucleotide sequence is represented in the public sequence databases by the accession NM_004354 and is described in this patent by Seq ID 322. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several
5 diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Cyclin G2 is repressed in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1F11 represents Hypothetical protein LOC51754. The protein sequence
10 encoded by Hypothetical protein LOC51754 is represented in the public databases by the accession XP_049657 and is described in this patent by Seq ID 323. The nucleotide sequence is represented in the public sequence databases by the accession AL137430 and is described in this patent by Seq ID 324. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic
15 products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Hypothetical protein LOC51754 is repressed in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In
20 a series of 5 patients with either ovarian or breast cancer, Hypothetical protein LOC51754 is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1F14 represents Butyrate response factor 1. The protein sequence encoded by Butyrate response factor 1 is represented in the public databases by the accession NP_004917 and is described in this patent by Seq ID 327. The nucleotide sequence is represented in the public sequence
25 databases by the accession NM_004926 and is described in this patent by Seq ID 328. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. HIF1alpha and EPAS1 are transcription factors which mediate the response to hypoxia of several genes, and have themselves been implicated in specific diseases. By adenoviral over-expression of EPAS1 we
30 show augmentation of the hypoxic induction of certain genes, further confirming their status as responsive to hypoxia. Butyrate response factor 1 has been shown to be induced by hypoxia to a greater degree following adenoviral over-expression of EPAS1. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer,

Butyrate response factor 1 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1F17 represents P8 protein (candidate of metastasis 1). The protein sequence encoded by P8 protein (candidate of metastasis 1) is represented in the public databases by the accession NP_036517 and is described in this patent by Seq ID 329. The nucleotide sequence is represented in the public sequence databases by the accession NM_012385 and is described in this patent by Seq ID 330. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. P8 protein (candidate of metastasis 1) is induced in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, P8 protein (candidate of metastasis 1) is up-regulated and also down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clones p1C1 and p1C2 represent CXCR4. The protein sequence encoded by CXCR4 is represented in the public databases by the accession NP_003458 and is described in this patent by Seq ID 331. The nucleotide sequence is represented in the public sequence databases by the accession NM_003467 and is described in this patent by Seq ID 332. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. CXCR4 is repressed in macrophages activated by LPS and gamma interferon. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. CXCR4 is induced in macrophages activated by TNFalpha. CXCR4 may act through the PI3-K pathway. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, CXCR4 is up-regulated and also down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1F3 represents Hypothetical protein XP_017131. The protein sequence encoded by Hypothetical protein XP_017131 is represented in the public databases by the accession XP_017131 and is described in this patent by Seq ID 333. The nucleotide sequence is represented in the public sequence databases by the accession XM_017131 and is described in this patent by Seq ID 334.

- 5 Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Hypothetical protein XP_017131 is repressed in macrophages activated by LPS and gamma
10 interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Hypothetical protein XP_017131 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

- The Oxford BioMedica clone p1F20 represents Proline-rich protein with nuclear targeting signal. The
15 protein sequence encoded by Proline-rich protein with nuclear targeting signal is represented in the public databases by the accession NP_006804 and is described in this patent by Seq ID 335. The nucleotide sequence is represented in the public sequence databases by the accession NM_006813 and is described in this patent by Seq ID 336. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic,
20 prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid
25 arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Proline-rich protein with nuclear targeting signal is induced in macrophages activated by TNFalpha. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Proline-rich protein with nuclear targeting signal is down-regulated in the
30 malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1F6 represents Hypothetical protein hqp0376. The protein sequence encoded by Hypothetical protein hqp0376 is represented in the public databases by the accession T08745 and is described in this patent by Seq ID 337. The nucleotide sequence is represented in the public sequence databases by the accession AF078844 and is described in this patent by Seq ID 338.

Hypothetical protein hqp0376 is a putative dead box protein as it shows high structural similarity to dead box proteins and yeast initiation factor 4A. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. HIF1alpha and EPAS1 are transcription factors which
5 mediate the response to hypoxia of several genes, and have themselves been implicated in specific diseases. By adenoviral over-expression of HIF1alpha we show augmentation of the hypoxic induction of certain genes, further confirming their status as responsive to hypoxia. Hypothetical protein hqp0376 has been shown to be induced by hypoxia to a greater degree following adenoviral over-expression of HIF1alpha. Hepatocytes are the main cell type of the liver and genes which are induced in response to
10 hypoxia in this cell type are relevant to development of diagnostics and therapeutics towards liver diseases involving hypoxia, including cirrhosis. Hypothetical protein hqp0376 is preferentially induced by hypoxia in hepatocytes. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are
15 especially relevant. Hypothetical protein hqp0376 is induced in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1F4 represents CYP1. The protein sequence encoded by CYP1 is represented in the public databases by the accession NP_000776 and is described in this patent by Seq ID 339. The nucleotide sequence is represented in the public sequence databases by the accession
20 NM_000785 and is described in this patent by Seq ID 340. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific
25 therapeutic products for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD. CYP1 is preferentially induced by hypoxia in monocytes or macrophages. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene
30 expression responses to both hypoxia and cytokines are especially relevant. CYP1 is induced in macrophages activated by LPS and gamma interferon. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have

utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. CYP1 is induced in macrophages activated by TNFalpha.

The Oxford BioMedica clone p1F15 represents SHB adaptor protein. The protein sequence encoded by SHB adaptor protein is represented in the public databases by the accession NP_003019 and is described
5 in this patent by Seq ID 341. The nucleotide sequence is represented in the public sequence databases by the accession NM_003028 and is described in this patent by Seq ID 342. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. SHB adaptor protein participates in tyrosine kinase-mediated signalling and the regulation of angiogenesis and apoptosis
10 [Dixelius J. 2000, Blood 95:3403-11]. Our surprising observation of the hypoxia regulation of this protein has clear medical utility in the diagnosis and treatment of vascular disease and cancer. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. SHB adaptor protein
15 is repressed in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1F13 represents Papillomavirus regulatory factor PRF-1. The protein sequence encoded by Papillomavirus regulatory factor PRF-1 is represented in the public databases by the accession NP_061130 and is described in this patent by Seq ID 343. The nucleotide sequence is represented in the public sequence databases by the accession AK023418 and is described in this patent
20 by Seq ID 344. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and
25 cytokines are especially relevant. Papillomavirus regulatory factor PRF-1 is repressed in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1A7 represents SLC31A2. The protein sequence encoded by SLC31A2 is represented in the public databases by the accession NP_001851 and is described in this patent by Seq ID 345. The nucleotide sequence is represented in the public sequence databases by the accession
30 NM_001860 and is described in this patent by Seq ID 346. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene

expression responses to both hypoxia and cytokines are especially relevant. SLC31A2 is induced in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1A21 represents UDP-glucose pyrophosphorylase 2. The protein sequence encoded by UDP-glucose pyrophosphorylase 2 is represented in the public databases by the accession
5 NP_006750 and is described in this patent by Seq ID 347. The nucleotide sequence is represented in the public sequence databases by the accession NM_006759 and is described in this patent by Seq ID 348. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

10 The Oxford BioMedica clones p1B4 and p1B5 represent Proline 4-hydroxylase, alpha polypeptide II. The protein sequence encoded by Proline 4-hydroxylase, alpha polypeptide II is represented in the public databases by the accession NP_004190 and is described in this patent by Seq ID 349. The nucleotide sequence is represented in the public sequence databases by the accession NM_004199 and is described in this patent by Seq ID 350. Hypoxia is an important feature of several diseases, and genes that respond
15 to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Proline 4-hydroxylase, alpha polypeptide II is repressed in macrophages
20 activated by LPS and gamma interferon and is also repressed in macrophages activated by IL-15. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Proline 4-hydroxylase, alpha polypeptide II is
25 induced in macrophages activated by TNFalpha. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Proline 4-hydroxylase, alpha polypeptide II is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clones p1B10, p1B11 and p1B12 represent Stearoyl-CoA desaturase. The protein
30 sequence encoded by Stearoyl-CoA desaturase is represented in the public databases by the accession NP_005054 and is described in this patent by Seq ID 351. The nucleotide sequence is represented in the public sequence databases by the accession NM_005063 and is described in this patent by Seq ID 352. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic

products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be
5 central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Stearoyl-CoA desaturase is induced in macrophages activated by TNFalpha.

The Oxford BioMedica clone p1B13 represents Diacylglycerol kinase, zeta. The protein sequence
10 encoded by Diacylglycerol kinase, zeta is represented in the public databases by the accession NP_003637 and is described in this patent by Seq ID 353. The nucleotide sequence is represented in the public sequence databases by the accession NM_003646 and is described in this patent by Seq ID 354. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic
15 products.

The Oxford BioMedica clone p1B22 represents Protease, serine, 11. The protein sequence encoded by Protease, serine, 11 is represented in the public databases by the accession NP_002766 and is described in this patent by Seq ID 355. The nucleotide sequence is represented in the public sequence databases by the accession NM_002775 and is described in this patent by Seq ID 356. Hypoxia is an important feature of
20 several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Protease, serine, 11 is
25 repressed in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Protease, serine, 11 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1B23 represents Interleukin 1 receptor antagonist. The protein sequence
30 encoded by Interleukin 1 receptor antagonist is represented in the public databases by the accession NP_000568 and is described in this patent by Seq ID 357. The nucleotide sequence is represented in the public sequence databases by the accession NM_000577 and is described in this patent by Seq ID 358. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic

products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia:

5 rheumatoid arthritis, atherosclerosis, cancer, COPD. Interleukin 1 receptor antagonist is preferentially induced by hypoxia in monocytes or macrophages. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions.

10 Interleukin 1 receptor antagonist is induced in macrophages activated by TNFalpha.

The Oxford BioMedica clone p1B24 represents NS1-binding protein. The protein sequence encoded by NS1-binding protein is represented in the public databases by the accession NP_006460 and is described in this patent by Seq ID 359. The nucleotide sequence is represented in the public sequence databases by the accession NM_006469 and is described in this patent by Seq ID 360. Hypoxia is an important feature

15 of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1C3 represents Activin A receptor, type I. The protein sequence encoded by Activin A receptor, type I is represented in the public databases by the accession NP_001096 and is described in this patent by Seq ID 361. The nucleotide sequence is represented in the public sequence

20 databases by the accession NM_001105 and is described in this patent by Seq ID 362. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Activin A is known to induce apoptosis [Hughes et al 1999, Prog Neurobiol 57:421-50], and so the regulation of its receptor by hypoxia has clear clinical significance. Hypoxia is frequently found in human tumours

25 where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Activin A receptor, type I is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1C4 represents FGF receptor activating protein 1. The protein sequence encoded by FGF receptor activating protein 1 is represented in the public databases by the accession

30 NP_055304 and is described in this patent by Seq ID 363. The nucleotide sequence is represented in the public sequence databases by the accession NM_014489 and is described in this patent by Seq ID 364. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. FGF has been shown to enhance survival of cardiac cells after ischaemic insult [Sheikh et al

2001, Am J Physiol Heart Circ Physiol 280:H1039-50], and so our observation of the hypoxia-regulation of the FGF receptor activating protein 1 is highly significant for the diagnosis and treatment of ischaemic disease. FGF receptor activating protein 1 is induced in macrophages treated with the inhibitory cytokine IL-10. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, FGF receptor activating protein 1 is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

- The Oxford BioMedica clone p1C5 represents Galectin 8. The protein sequence encoded by Galectin 8 is represented in the public databases by the accession NP_006490 and is described in this patent by Seq ID 365. The nucleotide sequence is represented in the public sequence databases by the accession
- 10 NM_006499 and is described in this patent by Seq ID 366. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Galectin 8 is an important tumour marker [for review see Bidon et al 2001, Int J Mol Med 8:245-50], and so its hypoxia-regulation is highly significant clinically.
- 15 The Oxford BioMedica clone p1C6 represents Glucose phosphate isomerase. The protein sequence encoded by Glucose phosphate isomerase is represented in the public databases by the accession NP_000166 and is described in this patent by Seq ID 367. The nucleotide sequence is represented in the public sequence databases by the accession NM_000175 and is described in this patent by Seq ID 368. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore
- 20 implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Glucose phosphate isomerase is induced in macrophages activated by IL-17 and also induced in
- 25 macrophages activated by IL-15. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Glucose phosphate isomerase is induced in macrophages activated by TNFalpha. Hypoxia is frequently found in human
- 30 tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Glucose phosphate isomerase is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1C7 represents D123. The protein sequence encoded by D123 is represented in the public databases by the accession NP_006014 and is described in this patent by Seq ID 369. The nucleotide sequence is represented in the public sequence databases by the accession NM_006023 and is described in this patent by Seq ID 370. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. D123 is repressed in macrophages activated by LPS and gamma interferon and is also repressed in macrophages activated by IL-15. D123 protein is an important regulator of the cell cycle [Onisto et al 1998, Exp Cell Res 242:451-9]. Recently it has been shown to be regulated by modification and turnover [Okuda et al 2001, Cell Struct Funct 26:205-14]. We have shown the hypoxia-regulation of this protein, and also of several prolyl hydroxylases which are known to target proteins for ubiquitination and proteasomal degradation. We believe that concerted hypoxic control of D123 and its regulating prolyl hydroxylase is part of the means of hypoxic regulation of cell growth and tissue re-modelling.

The Oxford BioMedica clone p1C8 represents DEC-1. The protein sequence encoded by DEC-1 is represented in the public databases by the accession NP_003661 and is described in this patent by Seq ID 371. The nucleotide sequence is represented in the public sequence databases by the accession NM_003670 and is described in this patent by Seq ID 372. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. DEC-1 is a helix-loop-helix transcription factor, and its hypoxia-regulation is highly significant. The response of renal epithelial cells to hypoxia is pertinent to kidney failure, especially regarding the medullary tissue. DEC-1 is preferentially induced by hypoxia in renal epithelial cells. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, DEC-1 is up-regulated and also down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1C9 represents RAB-8b protein. The protein sequence encoded by RAB-8b protein is represented in the public databases by the accession NP_057614 and is described in this patent by Seq ID 373. The nucleotide sequence is represented in the public sequence databases by the accession NM_016530 and is described in this patent by Seq ID 374. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. The hypoxia regulation of

this small GTP-ase, which is involved in intracellular membrane trafficking, is highly significant. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. RAB-8b protein
5 is induced in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1C10 represents Regulator of G-protein signalling 1. The protein sequence encoded by Regulator of G-protein signalling 1 is represented in the public databases by the accession NP_002913 and is described in this patent by Seq ID 375. The nucleotide sequence is represented in the public sequence databases by the accession NM_002922 and is described in this patent by Seq ID 376.

10 Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell
15 types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD. Regulator of G-protein signalling 1 is preferentially induced by hypoxia in monocytes or macrophages. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both
20 hypoxia and cytokines are especially relevant. Regulator of G-protein signalling 1 is repressed in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Regulator of G-protein signalling 1 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

25 The Oxford BioMedica clone p1C11 represents Polyubiquitin. The protein sequence encoded by Polyubiquitin is represented in the public databases by the accession BAA23632 and is described in this patent by Seq ID 377. The nucleotide sequence is represented in the public sequence databases by the accession AB009010 and is described in this patent by Seq ID 378. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and
30 have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Polyubiquitin is repressed in macrophages activated by LPS and gamma interferon. TNFalpha is an inflammatory cytokine, which acts

on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNF α therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Polyubiquitin is induced in macrophages activated by TNF α . Hypoxia is frequently found in human
5 tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Polyubiquitin is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1C12 represents Integrin, α 5. The protein sequence encoded by Integrin, α 5 is represented in the public databases by the accession NP_002196 and is described in
10 this patent by Seq ID 379. The nucleotide sequence is represented in the public sequence databases by the accession NM_002205 and is described in this patent by Seq ID 380. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Integrin, α 5 may play a rôle in the response to neuronal injury [King et al 2001, J Neurocytol 30:243-52]. Our observation of
15 hypoxia regulation of both COX-2 and integrin, α 5 supports the very recent suggestion that they may both function in recovery from cardiovascular injury [Hein et al 2001, Am J Physiol Heart Circ Physiol 281:H2378-84], which is pre-figured by our claims. Integrin, α 5 is induced by hypoxia in mammary epithelial cells, and may play an important role in cancer progression in that tissue through its function of regulating interaction with the extracellular matrix.

20 The Oxford BioMedica clone p1C13 represents Jk-recombination signal binding protein. The protein sequence encoded by Jk-recombination signal binding protein is represented in the public databases by the accession AAA60258 and is described in this patent by Seq ID 381. The nucleotide sequence is represented in the public sequence databases by the accession L07872 and is described in this patent by Seq ID 382. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus
25 are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Jk-recombination signal binding protein is repressed in macrophages activated by
30 LPS and gamma interferon. TNF α is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNF α therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Jk-recombination signal binding protein is induced in macrophages activated by TNF α . The important role of Jk-

recombination signal binding protein in the regulation of the immune response is thus modulated by hypoxia, and there are potentially many ways of exploiting that modulation in the design of diagnostics and therapeutics. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Jk-recombination signal binding
5 protein is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient. It is of particular interest and significance, in view of the escape from immunological surveillance of many tumours, that Jk-recombination signal binding protein is down-regulated.

The Oxford BioMedica clone p1C14 represents Abstrakt. The protein sequence encoded by Abstrakt is represented in the public databases by the accession NP_057306 and is described in this patent by Seq ID
10 383. The nucleotide sequence is represented in the public sequence databases by the accession NM_016222 and is described in this patent by Seq ID 384. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are
15 frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Abstrakt is repressed in macrophages activated by IL-15. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of
20 therapeutic, prognostic and diagnostic products for such inflammatory conditions. Abstrakt is induced in macrophages activated by TNFalpha. The general role of Abstrakt in the regulation of gene expression [Schmucker et al 2000, Mech Dev 91:189-96] implies particular significance to the recovery of cells from hypoxic insult.

The Oxford BioMedica clone p1C15 represents High-mobility group protein 2. The protein sequence
25 encoded by High-mobility group protein 2 is represented in the public databases by the accession NP_002120 and is described in this patent by Seq ID 385. The nucleotide sequence is represented in the public sequence databases by the accession NM_002129 and is described in this patent by Seq ID 386. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic
30 products.

The Oxford BioMedica clone p1C16 represents Decidual protein induced by progesterone. The protein sequence encoded by Decidual protein induced by progesterone is represented in the public databases by the accession NP_008952 and is described in this patent by Seq ID 387. The nucleotide sequence is represented in the public sequence databases by the accession NM_007021 and is described in this patent

- by Seq ID 388. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Decidual protein induced by progesterone is preferentially induced by hypoxia in mammary epithelial cells. Human decidual cells have not been tested, but we predict that
- 5 Decidual protein induced by progesterone is hypoxia-regulated in those cells. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Decidual protein induced by progesterone is repressed in macrophages activated by IL-17.
- 10 The Oxford BioMedica clone p1C19 represents GM2 ganglioside activator protein. The protein sequence encoded by GM2 ganglioside activator protein is represented in the public databases by the accession NP_000396 and is described in this patent by Seq ID 389. The nucleotide sequence is represented in the public sequence databases by the accession NM_000405 and is described in this patent by Seq ID 390. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore
- 15 implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia:
- 20 rheumatoid arthritis, atherosclerosis, cancer, COPD. GM2 ganglioside activator protein is preferentially induced by hypoxia in monocytes or macrophages. The hypoxia-inducibility of this protein in macrophages is likely to be clinically very significant. It is likely to play a role in the control of inflammation in asthma and inflammatory bowel disease, and in lipid metabolism and phosphatidylinositol-mediated signalling.
- 25 The Oxford BioMedica clone p1C20 represents CNOT8. The protein sequence encoded by CNOT8 is represented in the public databases by the accession NP_004770 and is described in this patent by Seq ID 391. The nucleotide sequence is represented in the public sequence databases by the accession NM_004779 and is described in this patent by Seq ID 392. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have
- 30 utility in the design of therapeutic, prognostic and diagnostic products.
- The protein sequence encoded by Similar to Nucleoside phosphorylase is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA430382 and is described in this patent by Seq ID 394. Hypoxia is an important

feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1P5 represents SCYA2. The protein sequence encoded by SCYA2 is represented in the public databases by the accession NP_002973 and is described in this patent by Seq ID 395. The nucleotide sequence is represented in the public sequence databases by the accession NM_002982 and is described in this patent by Seq ID 396. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. SCYA2 is induced in macrophages activated by LPS and gamma interferon and is also induced in macrophages activated by IL-17 or IL-15. Thus the role of SCYA2 in monocyte recruitment [Lu et al 1998, J Exp Med 187:601-8], which has clear relevance to the diagnosis and treatment of cardiovascular disease, cancer, rheumatoid arthritis, atherosclerosis and COPD, is enhanced by hypoxia. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. SCYA2 is repressed in macrophages activated by TNFalpha. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, SYCA2 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p2L23 represents Endothelin 1. The protein sequence encoded by Endothelin 1 is represented in the public databases by the accession NP_001946 and is described in this patent by Seq ID 397. The nucleotide sequence is represented in the public sequence databases by the accession NM_001955 and is described in this patent by Seq ID 398. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Endothelin 1 is induced in macrophages activated by LPS and gamma interferon. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have

utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. Endothelin 1 is induced in macrophages activated by TNF α . Endothelin 1 plays an important role in inducing proliferation of vascular smooth muscle cells. Its hypoxia-inducibility and thus its modulation to ameliorate the consequences of ischaemic insult, is of considerable clinical significance to the recovery
5 from injury, and angiogenesis.

The protein sequence encoded by Similar to Heat shock 70kD protein 4 is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA633656 and is described in this patent by Seq ID 400. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the
10 pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1K9 represents Lipocortin I. The protein sequence encoded by Lipocortin I is represented in the public databases by the accession NP_000691 and is described in this patent by Seq ID 401. The nucleotide sequence is represented in the public sequence databases by the accession NM_000700 and is described in this patent by Seq ID 402. Hypoxia is an important feature of several
15 diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Lipocortin I (also called annexin I) is an important anti-inflammatory mediator, and its hypoxia-inducibility has important implications for the diagnosis and treatment of ischaemic disease, cancer, atherosclerosis, and inflammatory diseases such as rheumatoid arthritis. Hypoxia is frequently found in human tumours where macrophage infiltrates are
20 also found. In a series of 5 patients with either ovarian or breast cancer, Lipocortin I is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Lipocortin I is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

25 The Oxford BioMedica clone p1K23 represents MYC. The protein sequence encoded by MYC is represented in the public databases by the accession NP_002458 and is described in this patent by Seq ID 403. The nucleotide sequence is represented in the public sequence databases by the accession NM_002467 and is described in this patent by Seq ID 404. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have
30 utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. MYC is repressed in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours

where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, MYC is up-regulated and also down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1K15 represents Alpha-2-macroglobulin. The protein sequence encoded
5 by Alpha-2-macroglobulin is represented in the public databases by the accession NP_000005 and is described in this patent by Seq ID 405. The nucleotide sequence is represented in the public sequence databases by the accession NM_000014 and is described in this patent by Seq ID 406. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is
10 a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD. Alpha-2-macroglobulin is preferentially induced by hypoxia in monocytes
15 or macrophages. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Alpha-2-macroglobulin is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1K8 represents SCYA4. The protein sequence encoded by SCYA4 is represented in the public databases by the accession XP_008449 and is described in this patent by Seq ID
20 407. The nucleotide sequence is represented in the public sequence databases by the accession XM_008449 and is described in this patent by Seq ID 408. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are
25 frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. SCYA4 is induced in macrophages activated by LPS and gamma interferon and is also induced in macrophages activated by IL-17. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in
30 expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. SCYA4 is induced in macrophages activated by TNFalpha. SCYA4 is a chemokine which is likely to be significant in inflammatory disease as a direct result of its hypoxic regulation. Hypoxia is frequently found in human tumours where macrophage

infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, SCYA4 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1M24 represents Sex hormone-binding globulin. The protein sequence encoded by Sex hormone-binding globulin is represented in the public databases by the accession
5 NP_001031 and is described in this patent by Seq ID 409. The nucleotide sequence is represented in the public sequence databases by the accession NM_001040 and is described in this patent by Seq ID 410. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

10 The Oxford BioMedica clone p1K7 represents ATP-binding cassette E1. The protein sequence encoded by ATP-binding cassette E1 is represented in the public databases by the accession NP_002931 and is described in this patent by Seq ID 411. The nucleotide sequence is represented in the public sequence databases by the accession NM_002940 and is described in this patent by Seq ID 412. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in
15 the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. ATP-binding cassette E1 is repressed in macrophages activated by LPS and gamma interferon.

20 The Oxford BioMedica clone p1K16 represents CCT6A. The protein sequence encoded by CCT6A is represented in the public databases by the accession NP_001753 and is described in this patent by Seq ID 413. The nucleotide sequence is represented in the public sequence databases by the accession NM_001762 and is described in this patent by Seq ID 414. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have
25 utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1K18 represents Colony-stimulating factor1. The protein sequence encoded by Colony-stimulating factor1 is represented in the public databases by the accession AAA52117 and is described in this patent by Seq ID 415. The nucleotide sequence is represented in the public sequence databases by the accession M37435 and is described in this patent by Seq ID 416. Hypoxia is an
30 important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease

sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Colony-stimulating factor1 is repressed in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1N1 represents GA17. The protein sequence encoded by GA17 is represented in the public databases by the accession NP_006351 and is described in this patent by Seq ID
5 417. The nucleotide sequence is represented in the public sequence databases by the accession NM_006360 and is described in this patent by Seq ID 418. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1K22 represents GPR44. The protein sequence encoded by GPR44 is
10 represented in the public databases by the accession NP_004769 and is described in this patent by Seq ID 419. The nucleotide sequence is represented in the public sequence databases by the accession NM_004778 and is described in this patent by Seq ID 420. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several
15 diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. GPR44 is repressed in macrophages activated by LPS and gamma interferon. GPR44 is most similar to the chemoattractant GPCR's [Marchese et al 1999, Genomics 1999 Feb 15;56(1):12-21]. Our demonstration of its hypoxic
20 regulation enables prediction of roles in diseases associated with transient hypoxia and macrophages. GPCR's are a druggable class of molecules, and represent an ideal route for pharmacological intervention.

The Oxford BioMedica clone p1K14 represents Keratin 6B. The protein sequence encoded by Keratin 6B is represented in the public databases by the accession NP_005546 and is described in this patent by Seq
25 ID 421. The nucleotide sequence is represented in the public sequence databases by the accession NM_005555 and is described in this patent by Seq ID 422. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Keratin 6B is induced in macrophages treated with the inhibitory cytokine IL-10. Keratin 6B is repressed in macrophages activated
30 by IL-17 and is also repressed in macrophages activated by IL-15.

The Oxford BioMedica clone p1K13 represents Lymphocyte adaptor protein. The protein sequence encoded by Lymphocyte adaptor protein is represented in the public databases by the accession NP_005466 and is described in this patent by Seq ID 423. The nucleotide sequence is represented in the

public sequence databases by the accession NM_005475 and is described in this patent by Seq ID 424. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

- 5 The Oxford BioMedica clone p1J20 represents Neuro-oncological ventral antigen 1. The protein sequence encoded by Neuro-oncological ventral antigen 1 is represented in the public databases by the accession NP_002506 and is described in this patent by Seq ID 425. The nucleotide sequence is represented in the public sequence databases by the accession NM_002515 and is described in this patent by Seq ID 426. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore
- 10 implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Neuro-oncological ventral antigen 1 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

- The Oxford BioMedica clone p1J22 represents Neutral sphingomyelinase (N-SMase) activation
- 15 associated factor. The protein sequence encoded by Neutral sphingomyelinase (N-SMase) activation associated factor is represented in the public databases by the accession NP_003571 and is described in this patent by Seq ID 427. The nucleotide sequence is represented in the public sequence databases by the accession NM_003580 and is described in this patent by Seq ID 428. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and
- 20 have utility in the design of therapeutic, prognostic and diagnostic products. Neutral sphingomyelinase (N-SMase) activation associated factor is induced in macrophages treated with the inhibitory cytokine IL-10. Neutral sphingomyelinase (N-SMase) activation associated factor is repressed in macrophages activated by IL-17 and is also repressed in macrophages activated by IL-15. We expect activation of of Neutral sphingomyelinase (N-SMase) to have an anti-inflammatory effect. This enzyme is known to
- 25 modulate the sphingomyelin second messenger cycle, potentially interacting with the oxidative system. Our demonstration of hypoxic regulation provides a crucial indication of the benefit of therapeutic intervention via sphingomyelinase (N-SMase) for the treatment of inflammatory diseases and diseases related to the hypoxic macrophage.

- The Oxford BioMedica clone p1K1 represents Cyclophilin F. The protein sequence encoded by
- 30 Cyclophilin F is represented in the public databases by the accession NP_005720 and is described in this patent by Seq ID 429. The nucleotide sequence is represented in the public sequence databases by the accession NM_005729 and is described in this patent by Seq ID 430. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found

in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Cyclophilin F is up-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1K3 represents Pleckstrin. The protein sequence encoded by Pleckstrin is represented in the public databases by the accession NP_002655 and is described in this patent by Seq ID 431. The nucleotide sequence is represented in the public sequence databases by the accession NM_002664 and is described in this patent by Seq ID 432. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD. Pleckstrin is preferentially induced by hypoxia in monocytes or macrophages. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Pleckstrin is induced in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Pleckstrin is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clones p1J19 and p1K2 represent CFFM4. The protein sequence encoded by CFFM4 is represented in the public databases by the accession NP_067024 and is described in this patent by Seq ID 433. The nucleotide sequence is represented in the public sequence databases by the accession NM_021201 and is described in this patent by Seq ID 434. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD. CFFM4 is preferentially induced by hypoxia in monocytes or macrophages. CFFM4 is induced in macrophages treated with the inhibitory cytokine IL-10. It has been suggested recently that CFFM4 is associated with mature cellular function in the monocytic lineage and that it may be a component of a

receptor complex involved in signal transduction [Gingras et al 2001, Immunogenetics 53:468-76]. Our demonstration of hypoxic-regulation opens possible routes of intervention in macrophage-related disease via this potentially important cell surface receptor. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, CFFM4 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1K5 represents Ribosomal protein L36a. The protein sequence encoded by Ribosomal protein L36a is represented in the public databases by the accession NP_000992 and is described in this patent by Seq ID 435. The nucleotide sequence is represented in the public sequence databases by the accession NM_001001 and is described in this patent by Seq ID 436. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1J17 represents SLC6A1. The protein sequence encoded by SLC6A1 is represented in the public databases by the accession NP_003033 and is described in this patent by Seq ID 437. The nucleotide sequence is represented in the public sequence databases by the accession NM_003042 and is described in this patent by Seq ID 438. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, SLC6A1 is up-regulated and also down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1J18 represents Synaptopodin. The protein sequence encoded by Synaptopodin is represented in the public databases by the accession NP_009217 and is described in this patent by Seq ID 439. The nucleotide sequence is represented in the public sequence databases by the accession NM_007286 and is described in this patent by Seq ID 440. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Synaptopodin is a component of the cytoskeleton which has particular importance in neurons, where it is involved in synaptic plasticity. Its hypoxia-regulation is clearly potentially significant in the context of neurological disease. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Synaptopodin is up-regulated and also down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1J15 represents TERA protein. The protein sequence encoded by TERA protein is represented in the public databases by the accession NP_067061 and is described in this patent by Seq ID 441. The nucleotide sequence is represented in the public sequence databases by the accession NM_021238 and is described in this patent by Seq ID 442. Hypoxia is an important feature of several
5 diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, TERA protein is up-regulated and also down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

10 The Oxford BioMedica clone p1K4 represents TSC-22. The protein sequence encoded by TSC-22 is represented in the public databases by the accession NP_006013 and is described in this patent by Seq ID 443. The nucleotide sequence is represented in the public sequence databases by the accession NM_006022 and is described in this patent by Seq ID 444. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have
15 utility in the design of therapeutic, prognostic and diagnostic products. TSC-22 is a transcriptional regulator of the leucine zipper class, and its hypoxic regulation is likely to have significant downstream effects which may be related to ischaemic disease. Thus it may provide important points of intervention in such diseases. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been
20 shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. TSC-22 is repressed in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, TSC-22 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

25 The Oxford BioMedica clone p2A14 represents an unannotated EST. The protein sequence encoded by this EST is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AA988110 and is described in this patent by Seq ID 446. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic,
30 prognostic and diagnostic products. The EST represented by Seq ID 446 is induced in macrophages treated with the inhibitory cytokine IL-10. The EST represented by Seq ID 446 is repressed in macrophages activated by IL-17 and is also repressed in macrophages activated by IL-15.

The Oxford BioMedica clone p1J23 represents Calgranulin A. The protein sequence encoded by Calgranulin A is represented in the public databases by the accession NP_002955 and is described in this

patent by Seq ID 447. The nucleotide sequence is represented in the public sequence databases by the accession NM_002964 and is described in this patent by Seq ID 448. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Calgranulin A, called by its
5 synonym S100A8, has been cited recently as "wound-regulated" [Thorey et al 2001, J Biol Chem 276:35818-25] which provides less precise support for our prior determination of its hypoxia-regulation. In its potential role, as a chemoattractant, it would be an important point of intervention for the modulation of inflammatory processes. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been
10 shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Calgranulin A is repressed in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Calgranulin A is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

15 The Oxford BioMedica clone p1J21 represents Replication factor C large subunit. The protein sequence encoded by Replication factor C large subunit is represented in the public databases by the accession NP_002904 and is described in this patent by Seq ID 449. The nucleotide sequence is represented in the public sequence databases by the accession NM_002913 and is described in this patent by Seq ID 450. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore
20 implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1J24 represents Signal recognition particle 19kD. The protein sequence encoded by Signal recognition particle 19kD is represented in the public databases by the accession NP_003126 and is described in this patent by Seq ID 451. The nucleotide sequence is represented in the
25 public sequence databases by the accession NM_003135 and is described in this patent by Seq ID 452. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1J16 represents cDNA: FLJ23019 fis, clone LNG00916. The protein
30 sequence encoded by cDNA: FLJ23019 fis, clone LNG00916 is not represented in the public databases by a protein accession. The nucleotide sequence is represented in the public sequence databases by the accession AK026672 and is described in this patent by Seq ID 454. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to

several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. The cDNA: FLJ23019 is, clone LNG00916 is induced in macrophages activated by LPS and gamma interferon and is also induced
5 in macrophages activated by IL-15.

The Oxford BioMedica clone p1J2 represents Proteasome subunit, alpha type, 4. The protein sequence encoded by Proteasome subunit, alpha type, 4 is represented in the public databases by the accession NP_002780 and is described in this patent by Seq ID 455. The nucleotide sequence is represented in the public sequence databases by the accession NM_002789 and is described in this patent by Seq ID 456.
10 Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1J9 represents MAFB. The protein sequence encoded by MAFB is represented in the public databases by the accession NP_005452 and is described in this patent by Seq ID
15 457. The nucleotide sequence is represented in the public sequence databases by the accession NM_005461 and is described in this patent by Seq ID 458. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. MAFB is a transcriptional regulator of the leucine zipper type, and is likely to play an important role in the mediation of the hypoxic
20 response, with attendant relevance to associated diseases.

The Oxford BioMedica clone p1J10 represents DNCL12. The protein sequence encoded by DNCL12 is represented in the public databases by the accession NP_006132 and is described in this patent by Seq ID 459. The nucleotide sequence is represented in the public sequence databases by the accession NM_006141 and is described in this patent by Seq ID 460. Hypoxia is an important feature of several
25 diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, gene X is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

30 The Oxford BioMedica clone p1J11 represents Chromobox homolog 3. The protein sequence encoded by Chromobox homolog 3 is represented in the public databases by the accession NP_057671 and is described in this patent by Seq ID 461. The nucleotide sequence is represented in the public sequence databases by the accession NM_016587 and is described in this patent by Seq ID 462. Hypoxia is an

important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1J5 represents SCYA7. The protein sequence encoded by SCYA7 is represented in the public databases by the accession NP_006264 and is described in this patent by Seq ID 5 463. The nucleotide sequence is represented in the public sequence databases by the accession NM_006273 and is described in this patent by Seq ID 464. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are 10 frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. SCYA7 is induced in macrophages activated by IL-15. SCYA7 is a chemoattractant protein which, considering its hypoxia-regulation, is likely to play an important role in inflammatory and ischaemic disease. TNFalpha is an inflammatory cytokine, which acts on macrophages, and has been shown to be central to the 15 pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. SCYA7 is repressed in macrophages activated by TNFalpha.

The Oxford BioMedica clone p1J11 represents Fatty-acid-Coenzyme A ligase, long-chain 2. The protein 20 sequence encoded by Fatty-acid-Coenzyme A ligase, long-chain 2 is represented in the public databases by the accession NP_066945 and is described in this patent by Seq ID 465. The nucleotide sequence is represented in the public sequence databases by the accession NM_021122 and is described in this patent by Seq ID 466. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, 25 prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Fatty-acid-Coenzyme A ligase, long-chain 2 is induced in macrophages activated by LPS and gamma interferon and also induced in macrophages activated by IL-17 or IL-15. 30 Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Fatty-acid-Coenzyme A ligase, long-chain 2 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1J8 represents Programmed cell death 5. The protein sequence encoded by Programmed cell death 5 is represented in the public databases by the accession NP_004699 and is

described in this patent by Seq ID 467. The nucleotide sequence is represented in the public sequence databases by the accession NM_004708 and is described in this patent by Seq ID 468. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

- 5 The Oxford BioMedica clone p1120 represents SCYA3L. The protein sequence encoded by SCYA3L is represented in the public databases by the accession CAA36397 and is described in this patent by Seq ID 469. The nucleotide sequence is represented in the public sequence databases by the accession X52149 and is described in this patent by Seq ID 470. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the
- 10 design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD. SCYA3L is
- 15 preferentially induced by hypoxia in monocytes or macrophages. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. SCYA3L is induced in macrophages activated by LPS and gamma interferon. TNFalpha is an inflammatory cytokine, which acts on
- 20 macrophages, and has been shown to be central to the pathophysiology and treatment of diseases including rheumatoid arthritis. Genes that change in expression in response to TNFalpha therefore have utility in the design of therapeutic, prognostic and diagnostic products for such inflammatory conditions. SCYA3L is induced in macrophages activated by TNFalpha.

- The Oxford BioMedica clone p113 represents Furin. The protein sequence encoded by Furin is
- 25 represented in the public databases by the accession NP_002560 and is described in this patent by Seq ID 471. The nucleotide sequence is represented in the public sequence databases by the accession NM_002569 and is described in this patent by Seq ID 472. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

- 30 The Oxford BioMedica clone p1112 represents Nuclear autoantigenic sperm protein. The protein sequence encoded by Nuclear autoantigenic sperm protein is represented in the public databases by the accession NP_002473 and is described in this patent by Seq ID 473. The nucleotide sequence is represented in the public sequence databases by the accession NM_002482 and is described in this patent by Seq ID 474. Hypoxia is an important feature of several diseases, and genes that respond to this

stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.

The Oxford BioMedica clone p1123 represents Ecotropic viral integration site 2A. The protein sequence encoded by Ecotropic viral integration site 2A is represented in the public databases by the accession
5 NP_055025 and is described in this patent by Seq ID 475. The nucleotide sequence is represented in the public sequence databases by the accession NM_014210 and is described in this patent by Seq ID 476. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. There is a prejudice in the art that the response to hypoxia is generic to all cell types. Contrary
10 to this, we show that genes are regulated by hypoxia to a greater degree in certain cell types, substantiating their utility in designing specific therapeutic products for diseases involving those cell types. Monocytes and macrophages have been implicated in the following diseases involving hypoxia: rheumatoid arthritis, atherosclerosis, cancer, COPD. Ecotropic viral integration site 2A is preferentially induced by hypoxia in monocytes or macrophages. Macrophages are key to several diseases involving
15 hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Ecotropic viral integration site 2A is repressed in macrophages activated by LPS and gamma interferon. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer,
20 Ecotropic viral integration site 2A is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p117 represents Sjogren syndrome antigen B. The protein sequence encoded by Sjogren syndrome antigen B is represented in the public databases by the accession
25 NP_003133 and is described in this patent by Seq ID 477. The nucleotide sequence is represented in the public sequence databases by the accession NM_003142 and is described in this patent by Seq ID 478. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Sjogren syndrome antigen B is preferentially induced by hypoxia in mammary epithelial cells. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In
30 these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. Sjogren syndrome antigen B is induced in macrophages activated by LPS and gamma interferon.

The Oxford BioMedica clone p1121 represents SCYA8. The protein sequence encoded by SCYA8 is represented in the public databases by the accession NP_005614 and is described in this patent by Seq ID

479. The nucleotide sequence is represented in the public sequence databases by the accession NM_005623 and is described in this patent by Seq ID 480. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several
5 diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. SCYA8 is induced in macrophages activated by LPS and gamma interferon and is also induced in macrophages activated by IL-15.
- 10 The Oxford BioMedica clone p119 represents GRO2. The protein sequence encoded by GRO2 is represented in the public databases by the accession NP_002080 and is described in this patent by Seq ID 481. The nucleotide sequence is represented in the public sequence databases by the accession NM_002089 and is described in this patent by Seq ID 482. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have
15 utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. GRO2 is induced in macrophages activated by LPS and gamma interferon and is also induced in macrophages activated by IL-
- 20 17. GRO2 encodes a chemokine which is likely to be involved in the inflammatory response. Its induction by hypoxia provides a potential route for intervention in diseases related to inflammation and ischaemia. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, GRO2 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.
- 25 The Oxford BioMedica clone p114 represents Small nuclear ribonucleoprotein D1. The protein sequence encoded by Small nuclear ribonucleoprotein D1 is represented in the public databases by the accession NP_008869 and is described in this patent by Seq ID 483. The nucleotide sequence is represented in the public sequence databases by the accession NM_006938 and is described in this patent by Seq ID 484. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore
30 implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products.
- The Oxford BioMedica clone p1124 represents GRO1. The protein sequence encoded by GRO1 is represented in the public databases by the accession NP_001502 and is described in this patent by Seq ID 485. The nucleotide sequence is represented in the public sequence databases by the accession

NM_001511 and is described in this patent by Seq ID 486. GRO1 has known chemotactic activity for neutrophils. GRO1 belongs to the intercrine alpha family of small CXC cytokines. GRO1 encodes a chemokine which is likely to be involved in the inflammatory response. Its induction by hypoxia provides a potential route for intervention in diseases related to inflammation and ischaemia. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Macrophages are key to several diseases involving hypoxia, and contribute to inflammatory processes. In these, macrophages are frequently activated by cytokines, which have been shown to be present at disease sites, so gene expression responses to both hypoxia and cytokines are especially relevant. GRO1 is induced in macrophages activated by LPS and gamma interferon and is also induced in macrophages activated by IL-17. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, GRO1 is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

The Oxford BioMedica clone p1118 represents Selectin L. The protein sequence encoded by Selectin L is represented in the public databases by the accession NP_000646 and is described in this patent by Seq ID 487. The nucleotide sequence is represented in the public sequence databases by the accession NM_000655 and is described in this patent by Seq ID 488. Hypoxia is an important feature of several diseases, and genes that respond to this stimulus are therefore implicated in the pathogenesis and have utility in the design of therapeutic, prognostic and diagnostic products. Selectin L shedding by leucocytes is one aspect of the induction of the inflammatory response. Hypoxic-regulation of Selectin L is clearly a significant factor in the induction of inflammation following ischaemic insult or in diseases in which transient ischaemic conditions occur. Modulation of this induction is one aspect of the present invention. Hypoxia is frequently found in human tumours where macrophage infiltrates are also found. In a series of 5 patients with either ovarian or breast cancer, Selectin L is down-regulated in the malignant tissue as compared to adjacent normal tissue in at least one patient.

TABLES**TABLE 1:** Hypoxia-inducible genes identified from clones only derived from the cardiomyoblast library

GENE NAME	SEQ ID		Accession
	protein	nucleotide	
Diacylglycerol kinase, zeta	353	354	NM_003646
CCR4 associated factor 1	391	392	AF053318
GM2 ganglioside activator protein.	389	390	X62078
Granulin	269	270	AK000607
Serine protease 11	355	356	Y07921
High mobility group 2 protein	385	386	M83665
Decidual protein induced by progesterone	387	388	NM_007021
DEAD-box protein abstrakt	383	384	NM_016222
IL-1 receptor antagonist	357	358	U65590
KIAA1376 protein	29	30	AB037797
Hypothetical protein KIAA0127	31	32	D50917
Hypothetical protein FLJ20308	33	34	AL137263
EST	91	92	AL390082
EST	89	90	AL117352
EST	77	78	AW664180

TABLE 2: Hypoxia-inducible genes identified from clones only derived from the macrophage libraries

GENE NAME	SEQ ID		Accession
	protein	nucleotide	
Metallothionein-2a	265	266	J00271
Metallothionein-1h	239	240	X64177
Metallothionein-1G	243	244	J03910
Interleukin 8	251	252	Y00787
Lactate dehydrogenase A	223	224	NM_005566
UDP-glucose pyrophosphorylase 2	347	348	NM_006759
Enolase 1	257	258	NM_001428
Enolase 2	273	274	NM_001975
Tissue factor / coagulation factor III / thromboplastin	225	226	NM_001993
proline 4-hydroxylase, alpha polypeptide I	231	232	NM_000917
proline 4-hydroxylase, alpha polypeptide II	349	350	NM_004199
NS1-binding protein	359	360	NM_006469
FGF receptor activating protein 1	363	364	AF159621
Adenylate kinase 3	263	264	NM_013410
Osteopontin	267	268	X13694
Aldolase C, fructose-bisphosphate	259	260	NM_005165
Galectin-8	365	366	AF193806
Regulator of G-protein signalling 1 (BL34)	375	376	S59049
Polyubiquitin UbC	377	378	AB009010
Activin A receptor type I	361	362	NM_001105
Glyceraldehyde-3-phosphate dehydrogenase	253	254	NM_002046
Phosphoglycerate kinase 1	255	256	NM_000291
Rab-8b	373	374	NM_016530
Glucose phosphate isomerase	367	368	NM_000175
D123 gene product (HT1080)	369	370	U27112
Integrin alpha 5	379	380	NM_002205
Triosephosphate isomerase 1	261	262	NM_000365
solute carrier family 31 (copper transporters),	345	346	NM_001860

member 2			
Jk-recombination signal binding protein	381	382	L07872
N-myc downstream regulated (NDRG1/ RTP)	229	230	D87953
Plasminogen activator inhibitor-1	235	236	M16006
Dec-1	371	372	NM_003670
FUSIN / CXCR4	331	332	NM_003467
Hypothetical protein FLJ20500	25	26	AK000507
DKFZP564D116 protein	27	28	AL050022
Hypothetical protein FLJ10134	23	24	AK000996
cDNA FLJ10433 fis NT2RP1000478	73	74	AK001295
ESTs	93	94	AW250104
ESTs	95	96	BE382614
ESTs	67	68	AW071063
ESTs	67	68	AW964331
ESTs	133	134	AA612751
Singleton EST (not in UniGene)	135	136	AI018611

The gene entitled "Jk-recombination signal binding protein" was found to be hypoxia-inducible using subtracted cDNA probes for hybridization, but with non-subtracted probes, where the hybridisation is quantitative, no signal was detected. This indicates that the gene is probably hypoxia-regulated but the absolute expression levels are very low.

TABLE 3: Hypoxia-inducible genes identified from clones derived from both macrophage and myoblast libraries.

GENE NAME	Accession	SEQ ID		Hypoxia/	Hypoxia/
		protein	nucleotide	normoxia	normoxia
				(macrophage)	(myoblast)
Solute carrier family 2, member 3	NM_006931	247	248	91.39	8.23
Solute carrier family 2, member 5	NM_003039	311	312	10.75	2.26
Adiphophilin	NM_001122	313	314	13.97	5.10
Hexokinase 2	NM_000189	249	250	11.50	6.25
Stearoyl-CoA desaturase	AB032261	351	352	3.74	2.31
cDNA DKFZp434O071	AF125392	75	76	2.31	2.75
Hypoxia-inducible protein 2	NM_013332	271	272	3.62	5.07

5 TABLE 4: Hypoxia responses amplified by HIF1alpha overexpression

Gene Name	Nucl Seq ID	Experimental Condition #								
		1	2	3	4	5	6	7	8	9
Metallothionein 2A	265	1	0.57	0.69	3.33	3.22	5.77	10.37	2.05	1.70
Metallothionein 1G	244	1	0.68	0.64	4.23	4.21	7.35	11.03	3.65	2.28
Hypothetical protein hqp0376	338	1	0.79	0.61	6.54	4.44	9.01	11.54	4.17	3.22
Novel Metallothionein	84	1	0.95	0.78	5.18	4.36	8.20	11.16	3.48	2.94

Legend: Data shown in the average of 4 repeat experiments. Experimental condition is as shown in the text. Values represent fold change as compared to untreated cells (condition 1).

TABLE 5: Hypoxia responses amplified by EPAS1 overexpression

Gene Name	Nucl	Experimental Condition #								
		Seq								
	ID	1	2	3	4	5	6	7	8	9
cDNA DKFZp586E1624	66	1	0.77	0.67	1.00	1.12	1.58	0.83	2.60	2.49
Butyrate response factor 1	328	1	0.74	0.64	1.60	1.64	1.57	1.23	2.19	3.20
hypothetical protein FLJ10134	24	1	0.62	0.53	2.73	2.09	2.80	2.87	4.20	3.65
EGL nine (C.elegans) homolog 3	86	1	1.34	0.81	1.98	1.90	2.02	1.94	2.81	3.12
ERO1 (S. cerevisiae)-like	68	1	1.02	1.30	4.26	4.14	4.76	4.12	4.91	6.44
hypothetical protein FLJ10134	24	1	0.68	0.53	2.03	1.97	3.01	2.46	3.67	2.95

Legend: Data shown is the average of 4 repeat experiments. Experimental condition is as shown in the text. Values represent fold change as compared to untreated cells (condition 1).

TABLE 6. Negative hypoxia responses amplified by HIF1alpha / EPAS1 overexpression

Gene Name	Nucl	Experimental Condition #								
		Seq								
	ID	1	2	3	4	5	6	7	8	9
Hypothetical protein CGI-117	48	1	0.83	0.87	0.42	0.42	0.32	0.34	0.33	0.27

Table 7: Genes induced by hypoxia (similar response +/- cell activation)

Row	TITLE	IMAGE Id	accession	SEQ ID		RATIO		
				protein	nucl	Hypoxia / Normoxia (resting)	Hypoxia / Normoxia (activated)	Activated / Resting (normoxia)
1	Activated leucocyte cell adhesion molecule	26617	R13558	277	278	1.46	1.86	0.46
2	MAX-interacting protein 1	435219	AA705886	279	280	2.55	3.18	n/d
3	BCL2/adenovirus E1B 19kD-interacting protein 3-like	814899	AA465697	217	218	2.50	3.48	0.41
4	Nuclear receptor co-repressor	488301	AA085748	281	282	1.38	1.75	0.65
5	Enolase 2, (gamma, neuronal)	789147	AA450189	273	274	2.87	4.98	1.32
6	Chitinase 3-like 2	47043	H10721	283	284	1.98	1.98	n/d
7	BACH1 transcription factor	2009495	A1336948	285	286	2.34	2.23	1.30
8	Solute carrier family 2, member 1	453589	AA679565	219	220	8.50	6.80	0.59
9	Phosphoglucomutase 1	843174	AA488504	287	288	1.43	1.83	n/d
10	PDGF beta	67654	T49539	221	222	1.66	1.64	1.09
11	PDGF beta	343320	W68169	221	222	1.86	1.67	0.84
12	CGI-109 protein	144862	R78570	289	290	1.42	1.94	n/d
13	SAP30	502142	AA126982	291	292	2.03	3.49	n/d
14	ATP-binding cassette transporter-1	827168	AA521292	293	294	2.04	2.24	1.20
15	SEC24 protein	712559	AA278134	295	296	2.87	3.97	n/d
16	Trinucleotide repeat containing 3	199367	R95691	297	298	1.92	1.38	1.35
17	Post-synaptic density protein 95	26021	R39954	299	300	1.79	1.64	1.63
18	Tumor protein D52	814306	AA459318	301	302	1.24	1.75	n/d
19	Cyclin-dependent kinase inhibitor p27kip1	854668	AA630082	303	304	2.36	1.57	2.19
20	phosphoinositide-3-kinase, catalytic, beta	506009	AA708437	305	306	1.44	2.11	0.26
21	cDNA FLJ13611 fis, clone PLACE1010802	49918	H15296	1	2	2.33	2.54	n/d

22	Solute carrier family 5, member 3	345743	W72666	307	308	3.33	4.42	1.14
23	PSCDBP	824531	AA490903	309	310	2.02	1.90	n/d
24	lactate dehydrogenase A	43550	H05914	223	224	2.13	2.23	1.23
25	Solute carrier family 2, member 5	190732	H38650	311	312	2.72	3.45	n/d
26	Adipophilin	435036	AA700054	313	314	6.28	2.39	n/d
27	Tissue factor	1928791	A1313387	225	226	1.34	2.21	0.62
28	Vascular endothelial growth factor	34778	R19956	227	228	1.53	1.97	n/d
29	RTP / NDRG1	842863	AA489261	229	230	3.40	3.06	2.38
30	Early development regulator 2	898328	AA598840	315	316	1.96	1.61	1.18
31	Procollagen-proline 4-hydroxylase alpha 1	838802	AA457671	231	232	2.69	2.32	1.31
32	B-cell translocation gene 1,	298268	N70463	317	318	1.91	2.08	1.84
33	SH3PX1	142139	R69163	319	320	1.81	1.15	1.87
34	Cyclin G2	823691	AA489752	321	322	1.70	2.47	n/d
35	BCL2/adenovirus E1B-interacting protein 3	783697	AA446839	233	234	4.37	6.52	n/d
36	BCL2/adenovirus E1B-interacting protein 3	359982	AA063521	233	234	3.09	5.00	n/d
37	NAG-5 protein	460618	AA700447	323	324	1.91	n/d	n/d
38	Cytochrome P450 IB1 (dioxin-inducible)	782760	AA448157	325	326	1.93	2.25	0.93
39	Plasminogen activator inhibitor, type 1	244307	N75719	235	236	1.81	2.78	n/d
40	Butyrate response factor 1	768299	AA424743	327	328	2.59	2.69	1.88
41	Butyrate response factor 1	413633	AA723035	327	328	2.35	2.36	1.76
42	p8 protein (candidate of metastasis 1)	80484	T64469	329	330	4.13	4.19	n/d
43	Fusin / CXCR 4	79629	T62491	331	332	2.16	1.97	1.07
44	solute carrier family 16, member 6	1638893	A1016779	333	334	1.96	n/d	n/d
45	solute carrier family 16, member 6	266389	N21654	333	334	2.00	n/d	n/d
46	Proline-rich protein with nuclear targeting signal (B4-2)	857002	AA669637	335	336	2.707	1.80	7.08

47	Cox-2	845477	AA644211	237	238	n/d	8.34	22.38
48	Glycogen synthase I (muscle)	45632	H08446	275	276	n/d	2.28	1.15
49	cDNA FLJ13700 fis, clone PLACE2000216, highly similar to SPECTRIN BETA CHAIN, BRAIN	261246	H98241	15	16	1.25	2.01	0.25
50	Hypothetical protein FLJ20037	142944	R71124	3	4	2.10	1.74	n/d
51	Hypothetical protein FLJ20037	451087	AA704517	3	4	2.36	1.73	n/d
52	hypothetical protein DKFZp434P0116	417863	W88781	5	6	2.18	1.85	0.95
53	KIAA0212	854874	AA630346	7	8	1.77	1.84	0.84
54	KIAA0914	283301	N51424	9	10	1.93	1.40	n/d
55	Hypothetical protein FLJ20281	244686	N54297	11	12	2.13	2.05	n/d
56	KIAA0876	809806	AA454753	13	14	2.80	1.82	n/d
57	DKFZP586G1122 protein	950778	AA608636	17	18	1.76	1.99	0.44
58	Putative zinc finger protein LOC55818	377452	AA055692	19	20	2.16	n/d	n/d
59	hypothetical protein PRO0823	194965	R88734	21	22	1.89	1.47	1.02
60	Hypothetical protein PRO0823	1486194	AA936866	21	22	2.17	1.31	0.46
61	cDNA DKFZp586H0324 clone DKFZp586H0324	130276	R21170	61	62	2.07	2.47	n/d
62	Clone 23785	376476	AA041362	63	64	2.50	2.29	n/d
63	Clone 23785	261834	H98855	63	64	2.14	1.97	0.49
64	cDNA DKFZp586E1624	284497	N52362	65	66	1.95	1.09	n/d
65	ESTs (UniGene annotated)	139558	R62339	79	80	1.73	1.75	1.69
66	ESTs (UniGene annotated)	897446	AA489477	81	82	1.74	1.83	n/d
67	ESTs (UniGene annotated)	126458	R06601	83	84	2.53	1.92	2.07
68	ESTs (UniGene annotated)	122982	R00332	85	86	2.43	3.96	n/d
69	ESTs (UniGene annotated)	811808	AA463469	87	88	1.79	1.31	1.24

70	ESTs(UniGene annotated)	203544	H56028	89	90	3.67	3.63	n/d
71	ESTs (UniGene annotated)	714437	AA293300	91	92	2.46	1.85	3.47
72	ESTs	810448	AA571116	67	68	4.87	2.97	1.22
73	ESTs	207275	H59618	97	98	2.61	1.24	1.04
74	ESTs	785928	AA449703	99	100	1.45	1.84	0.57
75	ESTs	827204	AA521311	101	102	1.80	1.55	1.27
76	ESTs	343695	W69170	103	104	1.78	1.48	1.69
77	ESTs	39145	R51835	105	106	1.49	1.72	1.04
78	ESTs	220608	H87770	107	108	1.47	2.09	1.09
79	ESTs	142087	R69248	109	110	1.57	1.70	n/d
80	ESTs	82171	T68844	111	112	2.44	2.16	1.19
81	ESTs	795325	AA454177	113	114	1.75	1.28	n/d
82	ESTs	366966	AA026562	115	116	1.27	1.70	n/d
83	ESTs	84419	T73780	117	118	1.43	2.22	1.07
84	ESTs	742611	AA401496	119	120	3.63	3.75	n/d
85	ESTs	277611	N49384	119	120	4.52	2.87	n/d
86	ESTs	823688	AA489636	121	122	2.11	n/d	n/d
87	ESTs	781311	AA446361	123	124	1.66	2.43	n/d
88	ESTs	1555201	AA931411	125	126	1.89	n/d	n/d
89	ESTs	131563	R24223	127	128	2.26	n/d	n/d
90	EST (singleton)	786657	AA451886	137	138	2.44	2.01	n/d
91	ESTs (ex-UniGene)	126393	R06520	139	140	1.59	1.81	0.86
92	ESTs(ex UniGene)	74054	T48278	141	142	1.92	1.05	n/d

Legend

5 The last 3 columns show mRNA expression as a ratio between the conditions being compared. Of these three columns the first two show expression in hypoxia relative to normoxia, done separately in resting macrophages or activated macrophages. The final column shows expression in activated macrophages relative to resting macrophages (both in normoxia) as a ratio. n/d = not determined due to low signal intensities. IMAGE ID and accession describe the exact identity of the arrayed clones and do not describe full length cDNA sequence database entries.

TABLE 8: Genes induced by hypoxia (greater response in resting cells)

Row	TITLE	IMAGE ID	accession	SEQ ID		RATIO		
				protein	nucl	Hypoxia / Normoxia (resting)	Hypoxia / Normoxia (activated)	Activated / Resting (normoxia)
1	Metallothionein 1H	214162	H77766	239	240	6.26	2.01	17.58
2	Metallothionein 1L	297392	N80129	241	242	18.55	2.21	7.57
3	metallothionein 1L	1899230	A1289110	241	242	5.89	1.70	9.63
4	Metallothionein-IG	202535	H53340	243	244	12.07	2.36	21.28
5	Metallothionein 1E (functional)	1472735	A A872383	245	246	10.16	2.04	4.66
6	RNA helicase-related metalothionein 1F	protein/245990	N55459	337	338	6.41	1.99	14.16
7	RNA helicase-related metalothionein 1F	protein/78353	T56281	337	338	5.19	1.54	12.00
8	Solute carrier family 2, member 3	753467	AA406551	247	248	7.67	4.69	4.78
9	Hexokinase 2	1637282	A1005515	249	250	7.32	3.27	0.62
10	DKFZp434E1723 DKFZp434E1723	clone 1593887	AA987423	69	70	2.04	1.34	1.49
11	Cytochrome P450, subfamily XXVIII, polypeptl	1761925	A1222585	339	340	2.16	0.72	2.92
12	Interleukin 8	549933	AA102526	251	252	5.65	0.86	382.80
13	SHB adaptor protein	768362	AA495786	341	342	1.87	0.84	0.39
14	ESTs	130835	R22252	129	130	1.97	0.93	0.83

Legend to Table 8

The last 3 columns show mRNA expression as a ratio between the conditions being compared. Of these three columns the first two show expression in hypoxia relative to normoxia, done separately in resting macrophages or activated macrophages. The final column shows expression in activated macrophages relative to resting macrophages (both in normoxia) as a ratio. n/d = not determined due to low signal intensities. IMAGE ID and accession describe the exact identity of the arrayed clones and do not describe full length cDNA sequence database entries.

TABLE 9: Genes induced by hypoxia (greater response in activated cells)

TITLE	SEQ ID				RATIO		
	IMAGE ID	accession	protein	nucleotide	Hypoxia / Normoxia (resting)	(activated)	Activated / Resting (normoxia)
Papillomavirus regulatory factor (PRF-1)	744983	AA625924	343	344	3.36	8.10	0.22
cDNA FLJ11041 fis, clone PLACE1004405	140301	R66924	71	72	1.46	3.19	2.50
ESTs (ex-UniGene)	139250	R68736	143	144	1.01	2.18	1.68

Legend

- 5 The last 3 columns show mRNA expression as a ratio between the conditions being compared. Of these three columns the first two show expression in hypoxia relative to normoxia, done separately in resting macrophages or activated macrophages. The final column shows expression in activated macrophages relative to resting macrophages (both in normoxia) as a ratio. n/d = not determined due to low signal intensities. IMAGE ID and accession describe the exact identity of the arrayed clones and do not describe full length cDNA sequence database entries.

TABLE 10: Genes repressed by hypoxia (greater response in activated cells)

row	TITLE	IMAGE ID	accession	SEQ ID		RATIO		
				protein	nucl	Hypoxia / Normoxia (resting)	Hypoxia / Normoxia (activated)	Activated/Resting (normoxia)
1	Mat-related leucine zipper homolog	77193	T50121	457	458	1.18	0.48	2.39
2	Alpha-2-macroglobulin	44180	H06516	405	406	1.11	0.54	1.98
3	KIAA0014	72527	AA292382	51	52	1.10	0.65	25.00
4	ESTs	178805	H49601	203	204	1.04	0.49	1.62
5	dynein, cytoplasmic, light intermediate polypeptide 2	1811870	AA454959	459	460	1.03	0.42	3.31
6	Heterochromatin-like protein 1	343490	W69106	461	462	1.01	0.60	3.21
7	Monocyte chemotactic protein 3	485989	AA040170	463	464	0.89	0.52	59.62
8	Fatty-acid-Coenzyme A ligase, long-chain 2	82734	T73556	465	466	0.88	0.52	6.85
9	Fatty-acid-Coenzyme A ligase, long-chain 2	2014138	A1361530	465	466	0.72	0.46	3.97
10	Programmed cell death 5 / TFAR19	502369	AA156940	467	468	0.78	0.59	1.57
11	cDNA FLJ14028 fis, clone HEMBA1003838	366156	AA062814	145	146	0.75	0.55	1.74
12	Small inducible cytokine A3	153355	R47893	469	470	0.69	0.29	8.74
13	Cytochrome c oxidase subunit VIc	42993	R59927	471	472	0.72	0.54	1.98
14	NASP histone-binding prot.	845415	AA644128	473	474	0.64	0.38	1.64
15	Hypothetical protein HSPC196	144902	R78498	53	54	0.63	0.48	1.62
16	Ecotropic viral integration site 2A	231675	H93149	475	476	0.63	0.35	1.51
17	Strogen syndrome antigen B	49970	H29484	477	478	0.53	0.32	7.86
18	Macrophage inflammatory protein 1b	205633	H62985	407	408	0.52	0.28	12.73
19	Monocyte chemotactic protein 1	768561	AA425102	395	396	0.46	0.11	213.89
20	Monocyte chemotactic protein 2	1911099	A1268937	479	480	undetectable	0.26	423.31
21	Endothelin 1	47359	H11003	397	398	undetectable	0.56	14.29

22	GRO2 /macrophage inflammatory protein 2a	153340	R50407	481	482	undetectable	0.66	12.16
23	Small nuclear ribonucleoprotein SM D1	47542	H16454	483	484	undetectable	0.22	11.01
24	Hypothetical protein FLJ11296	491460	AA150443	55	56	undetectable	0.51	8.96
25	GRO1 / macrophage inflammatory protein 2 precursor	324437	W46900	485	486	undetectable	0.48	15.29
26	GRO1 / macrophage inflammatory protein 2 precursor	323238	W42723	485	486	undetectable	0.40	7.92
27	Lymphocyte adhesion molecule 1	149910	H00756	487	488	undetectable	0.47	4.92
28	Sex hormone-binding globulin	82871	T69346	409	410	undetectable	0.36	3.57
29	ESTs	898045	AA598952	205	206	undetectable	0.53	2.37
30	Hypothetical protein bA395L14	842794	AA486203	57	58	undetectable	0.45	n/d

Legend to Table 10

5 The last 3 columns show mRNA expression as a ratio between the conditions being compared. Of these three columns the first two show expression in hypoxia relative to normoxia, done separately in resting macrophages or activated macrophages. The final column shows expression in activated macrophages relative to resting macrophages (both in normoxia) as a ratio. n/d = not determined due to low signal intensities. IMAGE ID and accession describe the exact identity of the arrayed clones and do not describe full length cDNA sequence database entries.

TABLE 11: Other genes repressed by hypoxia in macrophages

Row	TITLE	IMAGE ID	accession	PROTEIN	SEQ ID	RATIO		
						Hypoxia / Normoxia	(resting)	Activated/ Resting (normoxia)
1	Annexin A1	208718	H63077	401	402		0.86	1.27
2	ATP-binding cassette, sub-family (OABP), 1	E1593311	A1002355	411	412		0.62	0.79
3	ESTs	855583	AA664228	165	166		0.61	0.73
4	Chaperonin / Tcp zeta 1	45233	H07880	413	414		0.59	1.27
5	Chaperonin TCP1, subunit 6A zeta 1	45233	H07880	413	414		0.72	1.00
6	Colony stimulating factor 1 (macrophage)	173527	T55558	415	416		0.44	0.44
7	Colony stimulating factor 1 (macrophage)	11475574	AA878257	415	416		0.46	0.30
8	Dendritic cell protein (GA17)	563634	AA101348	417	418		0.59	0.97
9	G protein-coupled receptor 44	810403	AA464202	419	420		0.55	0.90
10	Heat shock 70kD protein 4	856567	AA633656	399	400		0.55	n/d
11	Keratin 6A	366481	AA026418	421	422		0.53	0.25
12	lymphocyte adaptor protein	294196	N71394	423	424		0.67	0.26
13	Neuro-oncological ventral antigen 1	2015354	A1362062	425	426		0.45	0.65
14	N-SMase / FAN	376644	AA046107	427	428		0.59	0.18
15	p67 myc protein	812965	AA464600	403	404		0.66	1.14
16	Peptidylprolyl isomerase F (cyclophilin F)	774726	AA442081	429	430		0.82	n/d
17	PLECKSTRIN	823779	AA490267	431	432		0.76	1.30
18	High affinity immunoglobulin epsilon receptor beta subunit	199185	R95749	433	434		0.65	1.22
19	High affinity immunoglobulin epsilon receptor beta subunit	79576	T62849	433	434		0.64	1.20

20	Ribosomal protein L44	884842	AA669359	435	436	0.73	0.50	1.36
21	Solute carrier family 6 No1	177967	H46254	437	438	0.62	0.54	1.38
22	Synaptopodin	178792	H49443	439	440	0.54	0.56	1.24
23	TERA protein	1521977	AA906997	441	442	0.44	0.34	0.46
24	TGF beta-stimulated protein TSC-22	868630	AA664389	443	444	0.56	0.72	0.83
25	Tubulin, beta, 2	1492104	AA88148	445	446	0.53	0.64	1.33
26	Calgranulin A	562729	AA086471	447	448	0.57	0.65	16.86
27	Replication factor C (145 KDa)	214537	H73714	449	450	0.66	0.50	2.80
28	Signal recognition particle 19 kD protein	754998	AA411407	451	452	0.63	0.49	2.77
29	Nucleoside phosphorylase	769890	AA430382	393	394	0.50	0.79	2.75
30	Transcription factor SUPT3H	1606865	AA996042	453	454	0.50	0.47	2.64
31	Proteasome component C9	399536	AA733040	455	456	0.64	0.58	2.09
32	Hypothetical nuclear factor SBB122	868119	AA634213	35	36	0.75	0.50	0.73
33	DKFZP434I116 protein	207379	H58884	37	38	0.52	0.43	1.04
34	Hypothetical prot. FLJ10206	487921	AA045286	39	40	0.61	1.00	0.24
35	hypothetical protein FLJ10815	1506046	AA905628	41	42	0.92	0.55	0.88
36	Hypothetical protein FLJ11100	811590	AA454607	43	44	0.56	0.47	1.17
37	hypothetical protein FLJ2064	868161	AA633831	45	46	0.81	0.47	0.85
38	Hypothetical protein HSPC111	825695	AA504814	47	48	0.43	0.44	1.05
39	hypothetical protein LOC51251	379941	AA778116	49	50	0.56	0.55	1.06
40	Hypothetical protein LOC51251	770997	AA427715	49	50	0.56	0.74	1.47
41	cDNA FLJ13016 fis, clone NT2RP3000624	77483	T58743	59	60	0.58	0.54	1.55
42	DKFZp564D016 (clone DKFZp564D016)	824665	AA482278	147	148	0.50	n/d	n/d
43	cDNA FLJ11302 fis, clone PLACE1009971	108351	T70612	149	150	0.75	0.49	1.36
44	NEDO FLJ10309 fis clone NT2RM200287	810026	AA455267	151	152	0.57	1.20	0.28

45	Sequence from clone RP11-39402 on ch 20	122147	T98503	153	154	0.51	0.99	n/d
46	ESTs (UniGene annotated)	731255	AA420992	155	156	0.67	0.53	0.86
47	ESTs (UniGene annotated)	434182	AA693797	157	158	0.46	0.37	0.77
48	ESTs (UniGene annotated)	788415	AA456437	159	160	0.50	0.52	0.89
49	ESTs (UniGene annotated)	49879	H28725	159	160	0.35	0.38	n/d
50	ESTs (UniGene annotated)	770954	AA429367	161	162	0.64	0.74	0.82
51	ESTs (UniGene annotated)	770935	AA434382	163	164	0.55	0.98	0.51
52	ESTs	34626	R44397	167	168	0.92	0.61	1.35
53	ESTs	1534589	AA923509	169	170	0.70	0.48	0.72
54	ESTs	417223	W87747	171	172	0.53	0.37	0.97
55	ESTs	854752	AA630167	215	216	0.64	0.54	0.87
56	ESTs	1569263	AA973568	173	174	0.52	0.45	0.62
57	ESTs	869440	AA679939	213	214	0.60	0.53	0.78
58	ESTs	123065	T98529	175	176	0.62	0.55	0.94
59	ESTs	364468	AA022679	177	178	0.59	0.54	1.30
60	ESTs	50635	H17921	179	180	0.45	0.43	0.94
61	ESTs	123858	R00766	181	182	0.50	0.50	0.75
62	ESTs	415195	W91958	183	184	0.54	0.59	1.28
63	ESTs	138865	R63694	185	186	0.56	0.66	1.32
64	ESTs	773308	AA425386	187	188	0.51	0.61	0.63
65	ESTs	1505857	AA909912	189	190	0.46	0.78	0.83
66	ESTs	122728	T99032	191	192	0.49	0.39	0.42
67	ESTs	202154	H52503	193	194	0.35	0.42	0.35
68	ESTs	502634	AA127017	195	196	0.52	0.88	0.54
69	ESTs	23005	R38647	197	198	0.51	n/d	n/d
70	ESTs	22500	T87233	199	200	0.55	n/d	n/d
71	ESTs	587398	AA130351	201	202	0.52	n/d	n/d

TABLE 12

	#1	#1	#2	#2	#3	#3	#4	#4	#5	#5	#6	#6	#7	#7	#8	#8	#9	#9	#10	#10	#11	#11												
	NO	HY	HY	NO	HY	NO	HY	NO	HY	NO	HY	NO	HY	NO	HY	NO	HY	NO	HY	NO	HY	HY												
	6hr	18hr	6hr	18hr	6hr	18hr	6hr	18hr	6hr	18hr	6hr	18hr	6hr	18hr	6hr	18hr	6hr	18hr	6hr	18hr	6hr	18hr												
Seq	ID	Clone																																
2	p1F12	1.26	0.83	1.12	2.59	2.41	2.84	1.36	1.42	1.75	0.41	0.35	0.41	0.70	1.59	0.48	0.46	0.71	0.41	0.45	0.26	0.42	0.70	0.65	1.55	1.34	1.68	0.80	1.06	0.95	1.21	1.68	1.38	
4	p1F2	1.39	2.37	1.74	1.18	1.78	1.62	0.74	0.67	0.50	0.80	1.12	1.24	1.30	1.98	0.94	1.38	1.29	0.83	1.55	0.94	1.42	1.92	4.03	0.73	0.57	0.87	0.31	0.33	0.42	0.69	1.03	0.79	
6	p1F10	1.41	1.07	0.83	1.57	1.71	1.71	1.60	1.92	1.67	0.79	0.66	0.56	0.69	1.48	1.17	0.94	1.50	0.98	2.39	0.54	0.59	0.77	0.87	1.01	0.88	0.72	0.77	0.91	0.76	1.86	1.90	1.62	
8	p1F19	0.54	0.53	0.78	0.99	1.04	1.25	1.18	1.23	1.22	0.44	0.54	0.72	0.37	0.29	0.94	0.67	0.71	2.69	4.29	1.99	0.57	0.87	1.12	1.21	1.79	1.07	1.16	1.57	1.56	1.06	0.98	0.93	
10	p1F8	0.63	1.34	1.17	0.87	1.04	1.04	0.53	1.07	1.09	0.50	1.20	1.96	0.32	0.99	0.74	1.21	1.69	0.60	2.33	0.81	1.22	2.39	3.16	0.91	0.91	1.82	0.33	1.03	1.98	1.33	2.23	3.62	
12	p1F5	1.41	1.04	0.98	1.12	1.63	1.35	2.09	2.40	1.91	0.78	0.65	0.88	0.81	1.67	0.83	0.75	1.46	0.57	1.93	0.41	0.96	1.83	2.23	0.85	0.86	0.75	0.63	0.81	0.81	2.06	2.83	1.93	
14	p1F18	1.34	0.69	1.02	1.43	1.65	1.90	1.98	1.62	2.17	0.74	0.55	0.72	0.40	0.57	0.74	0.70	0.90	0.31	0.40	0.27	0.78	1.21	1.00	1.33	1.41	1.00	1.16	1.67	1.54	1.04	1.01	1.04	
16	p1F7	0.69	0.35	0.75	1.63	1.35	1.67	1.99	5.27	2.95	1.05	0.91	0.71	0.47	0.47	0.05	0.08	0.05	0.91	0.86	0.52	0.01	0.03	0.02	1.87	1.16	1.17	3.87	5.09	5.00	1.13	1.59	1.52	
18	p1F21	0.68	0.58	0.50	0.85	0.85	0.74	0.80	0.66	0.52	0.61	0.66	0.63	0.27	0.59	2.16	5.39	5.25	3.00	5.81	5.25	2.78	5.02	8.14	1.90	2.11	0.96	0.87	1.43	1.29	0.92	1.34	1.01	
20	p1F9	0.36	0.47	0.86	0.95	2.47	3.09	3.60	2.70	3.01	0.43	0.93	1.02	0.23	0.54	0.29	0.62	0.81	0.66	2.81	2.93	0.30	1.24	1.98	1.26	3.69	5.69	0.93	2.81	4.90	0.76	1.30	1.41	
22	p1E13	1.76	1.30	0.95	1.04	1.78	1.01	0.59	0.66	0.55	0.89	0.65	0.72	1.04	1.50	3.73	3.36	3.44	0.77	0.91	0.52	4.94	5.66	7.44	0.80	0.76	0.83	0.48	0.48	0.53	1.25	1.37	1.45	
24	p1D1	0.88	1.98	5.49	1.47	2.03	3.59	1.32	2.69	2.05	0.63	2.45	5.51	0.94	1.79	0.10	0.33	0.79	0.84	1.45	1.75	0.06	0.08	0.15	0.18	0.21	2.61	0.41	1.22	2.15	0.52	1.53	2.90	
24	p1D2	1.03	0.99	3.89	1.76	3.50	4.79	0.73	1.87	1.78	1.07	2.22	4.16	0.66	1.01	0.13	0.23	0.36	0.45	0.69	1.35	0.14	0.14	0.20	0.91	1.82	5.00	0.71	2.33	2.12	1.15	2.37	4.15	
26	p1D4	0.41	1.88	1.95	0.83	1.80	2.19	1.57	1.80	1.54	0.42	1.42	1.50	0.49	2.36	0.31	0.39	0.66	4.30	30.5	28.2	0.07	0.88	1.59	1.59	0.17	0.22	2.68	0.96	4.76	10.5	0.28	1.01	0.69
28	p1D9	1.16	0.13	0.93	1.52	0.80	0.85	4.04	4.86	2.60	1.23	1.45	1.28	0.65	1.14	1.37	0.94	0.93	2.44	5.15	2.16	0.70	0.60	0.74	0.54	0.39	0.50	0.91	1.08	1.35	0.74	0.78	0.77	

30	p1D12	5.56	4.42	3.24	0.62	3.70	1.69	0.54	0.77	0.78	1.09	1.13	0.96	1.36	1.03	0.70	4.01	5.05	0.20	0.50	0.28	1.23	1.20	1.70	0.25	0.35	0.36	0.27	0.76	1.01	0.94	1.09	1.10
32	p1D15	0.81	0.54	1.48	1.10	1.39	1.05	0.74	0.86	0.69	0.77	0.90	0.81	0.38	0.73	0.54	0.65	1.10	1.18	1.38	1.14	0.33	0.44	0.53	1.73	1.72	1.69	2.13	4.23	8.70	1.24	1.31	1.89
34	p1D16	0.44	1.99	0.84	1.11	1.29	1.22	2.07	2.43	1.61	0.42	0.78	1.32	0.61	1.18	0.66	0.98	1.42	1.67	4.88	1.93	0.70	0.86	1.12	0.33	0.24	2.10	0.73	1.20	1.81	0.48	0.88	1.28
36	p1J13	1.15	0.62	0.54	1.24	0.91	0.84	1.94	1.73	1.66	1.08	0.75	0.50	0.66	0.69	0.51	0.39	0.46	1.61	2.09	0.97	0.58	0.55	0.60	2.18	1.53	0.88	1.41	1.42	1.10	1.49	1.21	0.96
38	p1I22	0.72	0.65	1.20	1.97	1.77	1.49	1.57	1.60	1.75	0.95	0.90	0.86	0.95	1.59	1.10	0.74	0.82	0.58	0.59	0.16	0.48	1.16	0.82	1.45	0.97	0.85	0.87	0.96	1.00	2.15	1.26	1.26
40	p1J6	0.91	0.96	1.04	1.08	0.99	0.96	0.97	1.21	0.87	0.99	1.04	1.07	1.18	1.60	1.10	1.29	1.48	0.60	0.33	0.39	1.79	1.70	1.50	0.88	1.23	0.75	0.63	0.48	0.75	0.98	0.93	0.99
42	p1I5	3.02	1.24	2.56	1.56	3.17	1.47	1.46	1.76	1.21	1.16	1.32	0.93	0.54	0.88	0.73	1.07	0.86	0.52	1.00	0.31	0.35	0.56	0.55	0.90	1.09	0.56	0.81	0.65	0.38	1.71	2.18	1.36
44	p1I13	8.65	6.65	3.11	0.71	3.20	0.89	0.96	1.11	0.84	1.92	1.66	1.08	1.25	0.71	2.09	2.43	3.69	0.66	0.22	0.09	0.80	0.86	0.71	1.71	0.85	0.82	1.23	0.92	0.65	0.77	0.81	0.53
46	p1I17	0.88	0.55	0.63	2.03	1.13	1.38	1.93	4.01	2.52	1.16	0.92	0.88	0.23	0.32	0.67	0.51	0.64	0.74	0.71	0.42	0.54	0.75	0.73	1.67	1.42	2.09	1.44	1.48	1.22	2.14	2.56	1.78
48	p1I15	2.06	1.31	0.49	1.19	0.53	0.43	1.88	1.66	0.97	1.47	0.74	0.91	1.24	1.66	0.61	0.34	0.31	2.06	1.98	0.44	0.29	0.24	0.22	2.30	1.31	0.96	1.57	0.94	0.41	0.94	0.80	0.45
50	p1I17	0.79	0.56	0.64	1.20	0.69	0.80	2.13	3.28	1.74	0.91	0.96	0.99	0.47	0.65	0.57	0.34	0.37	1.86	1.54	1.42	0.28	0.36	0.49	0.38	2.22	2.73	2.24	2.29	2.35	1.19	0.94	0.91
54	p1I14	1.26	0.51	0.35	1.21	0.95	0.47	1.04	1.45	0.93	0.90	0.43	0.29	0.73	0.82	1.40	1.18	1.05	1.83	1.14	0.40	0.74	0.78	0.93	2.06	1.20	0.52	1.57	0.90	0.39	1.25	1.21	1.06
56	p1I18	4.53	3.14	2.83	1.51	2.71	1.50	0.89	1.04	0.76	1.24	1.09	1.00	0.65	1.04	0.93	1.26	1.50	0.31	0.27	0.27	0.46	0.50	0.59	0.90	1.14	0.84	0.36	0.50	0.56	1.24	1.76	1.35
58	p1I16	0.70	0.51	0.41	2.31	1.46	1.60	3.37	2.91	2.73	0.51	0.51	0.43	0.76	0.64	1.03	0.82	0.99	10.6	17.9	4.96	0.51	0.73	0.42	0.83	0.79	0.42	2.34	3.52	2.38	1.54	1.61	1.06
60	p1I11	0.63	0.62	0.34	1.64	1.36	1.19	2.75	1.62	1.38	0.63	0.37	0.40	0.59	0.84	0.56	0.51	0.57	1.13	1.47	0.49	0.64	0.61	0.55	3.68	2.80	2.80	1.54	1.51	1.07	1.89	1.81	1.62

62	p1E8	0.67	0.47	0.71	1.41	1.14	1.22	1.59	3.29	1.99	1.08	1.07	0.95	0.25	0.40	0.52	0.48	0.66	1.23	1.31	0.97	0.41	0.63	0.57	1.93	1.25	1.44	1.50	2.35	2.34	0.98	1.19	0.98
64	p1E18	1.26	1.07	0.72	1.45	2.07	1.10	0.77	0.93	0.73	0.93	0.71	0.80	0.57	1.23	6.36	5.38	5.48	0.46	0.48	0.34	2.58	2.30	3.06	1.14	0.78	0.72	0.51	0.62	0.68	1.56	1.91	1.78
66	p1E16	0.84	0.56	1.06	0.96	0.95	0.88	1.90	6.76	5.00	2.43	2.67	2.25	0.32	0.49	0.62	0.81	1.01	3.17	3.38	1.75	0.16	0.53	0.90	0.18	0.22	0.17	0.75	1.05	1.15	1.25	1.08	1.31
68	p1D5	1.71	1.21	1.40	1.58	1.81	1.66	2.13	1.93	2.05	0.84	0.79	0.81	0.78	0.90	0.75	0.80	0.98	0.47	0.82	0.51	0.52	0.74	0.71	1.13	0.91	1.23	1.18	1.58	1.21	1.05	1.24	1.13
68	p1D6	0.24	1.42	2.51	0.28	0.61	0.93	1.77	1.57	1.85	0.25	1.35	2.82	0.34	0.52	0.61	1.49	2.56	1.76	11.9	9.56	0.51	1.35	2.15	0.10	0.13	1.08	0.39	1.28	3.59	0.26	0.39	1.00
70	p1E12	0.54	0.41	0.62	0.98	1.29	0.67	0.98	1.47	1.13	0.83	0.84	0.87	1.11	1.17	0.59	1.01	1.03	0.67	1.66	0.89	0.53	0.88	1.07	1.27	1.41	1.18	1.21	1.59	1.86	1.11	1.20	1.06
72	p1E10	1.09	0.63	1.03	4.31	3.17	5.08	2.27	4.17	3.76	1.03	1.01	1.06	0.75	0.74	0.74	0.79	0.91	1.62	1.31	0.79	0.60	0.78	0.79	0.74	0.70	0.48	2.12	0.88	1.20	1.37	1.53	1.27
74	p1C21	3.92	2.75	2.31	0.93	2.33	0.92	1.81	2.27	1.45	0.91	0.96	0.95	0.41	0.33	2.10	2.68	2.86	0.60	0.66	0.49	0.60	0.65	0.85	0.41	0.49	0.15	1.16	1.65	0.93	1.09	1.12	1.00
76	p1D10	0.45	0.88	1.49	0.60	1.80	1.63	1.34	2.61	2.82	0.34	1.48	1.83	0.65	1.21	0.36	0.49	0.89	1.45	2.68	2.94	0.35	0.35	0.53	0.75	1.64	3.48	0.80	4.59	6.54	0.50	1.17	1.22
78	p1D13	2.93	1.83	2.94	0.80	3.69	1.10	1.51	3.28	3.02	0.89	1.59	1.30	0.67	0.64	0.50	0.75	0.82	0.50	0.76	0.45	0.15	0.27	0.42	0.52	2.31	2.24	1.08	3.73	2.25	0.92	2.02	2.80
80	p1E9	1.54	0.89	1.48	1.18	1.80	1.53	0.68	0.82	0.60	0.60	0.64	0.73	0.92	1.62	4.20	3.97	5.19	0.25	0.42	0.18	2.33	3.63	5.01	0.93	0.80	0.85	0.57	0.61	0.61	1.58	1.37	1.17
82	p1F1	0.58	0.33	0.52	1.42	1.19	1.14	1.64	1.70	1.49	0.95	0.63	0.54	0.37	0.50	0.62	0.61	0.51	0.91	0.90	0.30	0.34	0.53	0.66	1.66	1.23	1.06	1.46	1.58	1.03	1.25	1.24	1.06
84	p1E7	1.20	1.29	1.63	0.12	0.18	0.22	1.40	1.03	0.71	2.36	2.17	2.85	4.89	6.75	0.08	1.07	0.86	0.87	2.06	1.32	0.03	0.75	1.20	0.04	0.05	0.04	3.09	2.60	3.60	0.60	0.70	0.74
86	p1E6	1.36	1.09	1.01	0.79	0.92	1.04	0.79	1.35	1.04	1.32	1.87	1.63	3.25	5.97	0.34	0.74	0.85	0.58	2.26	2.67	0.25	0.57	0.65	0.46	0.52	0.82	1.89	1.89	3.91	1.17	1.42	1.57
88	p2B1	1.51	0.97	0.80	2.39	2.31	2.25	1.06	1.22	1.18	1.10	0.77	0.51	0.72	1.19	0.60	0.50	0.77	0.97	1.30	0.60	0.65	0.87	0.56	44.4	43.0	40.8	1.00	0.82	0.90	2.60	2.73	2.05

90	p1D14	0.67	0.82	1.64	0.60	3.42	1.80	1.78	5.24	3.57	1.07	2.33	2.11	0.68	0.67	0.29	0.74	0.52	0.58	1.23	1.29	0.24	0.75	0.72	3.09	9.75	4.95	2.47	5.03	3.55	0.79	2.49	2.20
92	p1D17	5.32	5.30	3.77	0.46	3.08	1.47	0.41	0.60	0.43	1.05	0.92	0.91	1.17	1.13	3.73	3.80	5.77	0.58	1.01	0.76	0.98	1.26	1.31	0.35	0.56	0.65	0.79	2.26	4.40	0.83	1.71	0.82
92	p1P14	0.43	0.66	0.74	1.22	1.60	1.14	1.00	1.53	1.32	0.49	0.73	1.15	0.43	0.69	0.21	0.45	0.88	2.63	11.8	5.90	0.14	0.71	0.84	1.46	1.93	3.86	4.65	15.4	34.0	0.82	2.07	1.12
94	p1C24	0.92	0.88	0.53	1.76	1.65	1.37	1.08	1.09	0.89	1.23	1.07	0.85	0.68	1.28	0.60	0.77	3.67	1.02	2.70	1.61	1.06	1.66	0.88	1.22	0.95	1.30	0.77	0.67	0.81	1.31	1.62	1.57
96	p1D3	0.62	1.56	1.13	1.83	0.85	0.73	0.09	4.01	3.43	0.29	0.52	0.71	0.15	0.28	5.72	5.67	5.80	5.79	9.49	5.80	3.32	3.26	5.79	0.36	0.29	0.54	0.38	0.42	0.77	0.67	0.93	0.98
98	p1E14	1.09	1.92	0.89	1.93	1.46	0.96	1.23	1.58	0.96	0.88	0.71	1.26	0.87	1.68	0.67	0.89	2.54	1.20	9.33	2.55	0.38	0.74	1.03	0.80	0.74	0.82	0.76	0.94	0.99	1.32	1.63	1.27
100	p1E20	0.63	0.61	1.12	1.47	1.34	0.99	1.45	1.94	1.49	0.68	0.88	0.98	0.19	0.34	0.36	0.43	0.37	1.33	1.44	0.65	0.23	0.43	0.33	0.97	1.30	1.16	0.94	1.89	1.78	1.12	1.51	1.34
102	p2A24	1.21	0.65	0.67	2.42	2.21	2.48	2.35	3.05	2.59	1.10	0.76	0.68	0.68	0.91	0.28	0.22	0.38	0.37	0.40	0.20	0.18	0.26	0.17	1.73	1.59	1.32	1.60	1.85	1.36	1.95	1.44	
104	p1E17	0.94	0.56	0.99	0.95	1.55	0.96	1.93	2.41	2.37	0.89	0.88	0.80	0.53	0.73	0.84	1.15	0.90	0.29	0.39	0.24	1.96	2.03	2.06	1.27	1.47	0.91	1.13	1.31	1.33	0.97	1.19	1.01
106	p1E19	0.79	0.65	0.62	2.64	2.37	2.99	2.10	1.61	1.77	0.61	0.40	0.64	0.42	0.89	0.68	0.76	1.08	0.41	0.50	0.32	0.38	0.90	0.47	1.30	1.11	1.17	1.29	1.52	1.39	1.14	1.69	1.39
108	p1E15	2.82	1.46	2.04	1.42	2.65	1.00	0.92	1.10	0.78	1.83	2.00	1.73	0.82	0.72	1.21	1.35	1.27	1.12	1.01	0.85	0.52	0.65	1.11	0.44	0.56	0.30	0.28	0.29	0.32	0.93	1.05	0.75
110	p1E11	1.36	1.12	0.87	2.09	2.86	3.03	0.74	0.91	0.88	0.89	0.59	0.79	0.77	2.09	2.76	2.97	2.90	0.63	0.84	0.34	0.64	1.47	1.35	1.18	1.16	0.64	0.74	0.79	0.88	1.79	2.49	1.93
112	p1E23	1.19	0.65	1.55	1.24	2.11	1.42	0.75	1.15	1.02	0.86	1.09	1.09	0.35	0.40	0.31	0.36	0.48	0.88	1.02	0.73	0.30	0.55	0.66	3.17	2.93	2.98	1.21	2.36	2.53	2.23	2.61	2.23
114	p1E21	0.82	0.52	0.85	2.08	2.28	2.04	2.14	2.04	2.27	0.74	0.49	0.79	0.51	1.04	0.48	0.67	0.67	0.35	0.34	0.21	0.56	1.14	0.89	1.52	1.62	1.30	1.42	1.69	1.38	1.02	1.32	1.03
116	p1D23	0.95	0.47	0.68	1.21	1.62	1.30	0.99	1.17	1.10	1.39	1.20	1.05	0.35	0.49	0.28	0.27	0.34	0.80	0.85	0.66	0.39	0.49	0.53	2.49	2.28	1.52	1.02	1.77	1.19	1.25	1.21	1.17

118	p1D24	0.98	0.71	1.39	1.55	1.98	2.16	1.18	1.47	1.01	0.74	0.49	0.67	1.53	1.94	0.90	0.98	0.96	0.71	0.79	0.49	0.53	0.90	0.93	0.81	0.51	0.43	4.63	10.9	11.79	1.85	2.23	1.65
120	p1D22	0.70	0.89	1.98	0.89	2.66	2.13	0.71	1.40	1.45	0.72	1.61	2.10	0.26	0.65	0.22	0.96	0.84	0.92	2.62	2.82	0.32	1.26	2.00	2.03	7.51	5.60	0.91	3.98	5.11	0.62	2.09	2.39
122	p1E2	3.39	1.91	1.92	1.56	1.78	1.65	1.11	1.09	1.03	3.44	3.32	1.96	0.68	0.78	0.76	0.73	0.68	0.51	0.58	0.32	0.53	0.51	0.68	3.13	2.02	1.22	1.15	1.93	1.52	0.93	0.61	0.59
124	p1E1	2.91	1.31	2.14	3.07	6.21	2.25	1.08	2.11	1.28	1.02	1.01	1.21	0.56	0.80	0.64	1.16	1.25	0.29	0.42	0.29	0.28	0.47	0.91	1.03	1.24	0.88	0.32	0.61	0.60	3.03	4.39	2.33
126	p1E4	0.65	0.51	0.79	1.06	1.24	1.18	1.59	1.57	1.70	0.80	0.76	0.91	0.34	0.39	0.97	1.13	1.21	1.65	1.17	0.82	0.74	1.28	1.71	0.95	0.97	0.93	0.78	1.35	1.44	1.30	1.48	1.35
128	p1D18	8.78	5.01	4.93	1.12	5.44	2.18	0.58	0.80	0.50	1.25	1.25	0.99	1.15	1.18	1.34	1.93	2.33	0.61	1.12	0.59	0.21	0.47	0.44	0.47	0.56	0.34	0.81	0.66	0.93	1.67	2.06	1.43
130	p1D21	1.56	0.98	0.96	1.18	2.30	1.47	0.48	0.78	0.60	0.67	0.60	0.59	0.40	0.63	0.91	1.44	1.58	0.75	0.63	0.26	0.63	0.94	1.35	1.42	1.59	1.61	4.36	8.21	7.50	0.98	1.29	1.38
132	p1C22	3.70	2.08	2.77	0.94	2.43	1.34	0.92	1.26	0.74	1.85	1.25	1.49	0.91	0.77	2.36	2.86	3.39	0.24	0.24	0.21	1.34	1.44	2.22	0.54	0.74	0.36	0.34	0.38	0.45	0.93	1.16	0.89
134	p1C23	0.91	0.93	0.88	1.84	1.17	1.36	3.92	3.52	2.77	1.02	1.07	1.63	0.60	0.73	0.67	0.54	0.85	0.87	1.53	0.69	0.31	0.63	0.59	1.02	0.68	1.28	1.64	2.00	2.74	1.27	1.80	1.29
136	p1D11	0.54	1.09	1.35	0.73	1.57	1.74	2.78	4.49	3.68	0.60	1.22	1.63	0.80	1.83	0.22	0.74	0.90	0.33	2.31	1.23	0.26	1.00	1.34	0.38	0.41	1.04	0.49	2.07	2.79	0.86	1.77	1.60
138	p1E3	2.55	0.54	0.61	6.62	1.75	1.81	0.34	0.15	0.09	1.15	0.42	0.25	0.28	0.41	2.49	2.94	3.29	0.99	0.43	0.18	5.87	5.74	7.79	0.35	0.44	0.17	1.79	0.50	0.31	2.08	1.61	1.29
140	p1D20	0.75	0.93	1.16	1.35	1.72	1.18	1.63	2.62	2.78	0.57	0.54	0.75	0.43	0.69	0.66	0.77	0.78	1.34	1.76	0.65	0.34	0.71	0.72	2.43	2.48	1.46	0.66	1.80	1.75	3.27	3.69	4.17
142	p1E5	0.86	0.56	0.60	1.59	1.42	1.41	0.94	1.32	0.94	0.41	0.82	0.56	5.88	11.0	2.01	5.20	3.83	0.83	1.03	0.57	0.83	1.57	1.37	0.67	0.70	0.68	0.55	0.64	0.78	1.19	1.55	1.19
144	p1D19	3.52	1.95	2.39	1.10	3.29	1.54	0.96	1.23	0.82	1.87	1.69	1.80	0.59	0.66	1.50	2.04	1.97	0.16	0.12	0.14	0.80	1.04	1.08	0.46	0.58	0.25	0.49	0.46	0.59	1.34	1.62	1.22
146	p2A15	0.89	0.64	0.65	2.04	1.91	1.67	2.70	3.55	3.33	1.25	1.20	1.00	0.74	0.69	0.44	0.50	0.54	0.64	0.82	0.41	0.64	0.39	0.40	1.61	1.48	1.44	2.08	2.36	2.45	1.01	1.17	1.02

148	p1H14	0.96	0.54	0.67	1.30	0.68	1.11	2.36	2.45	1.69	1.10	0.90	0.47	0.20	0.24	0.38	0.26	0.26	2.20	3.06	1.28	0.15	0.27	0.19	1.67	1.28	0.67	1.87	2.43	1.77	1.06	1.00	0.68
150	p1I2	1.89	1.15	1.70	2.50	3.12	2.26	1.57	1.92	1.47	0.73	0.60	0.85	0.46	0.66	0.95	0.63	0.93	1.17	1.93	0.88	0.16	0.30	0.31	1.14	1.13	1.01	0.36	0.55	0.54	0.57	5.13	2.75
152	p1H12	2.00	1.52	1.68	0.91	1.02	1.40	0.66	0.88	0.62	1.81	1.36	1.48	1.96	2.49	2.13	2.37	2.71	0.54	0.27	0.32	2.90	3.25	2.42	0.62	0.85	0.44	0.36	0.34	0.40	0.87	1.00	0.87
154	p1I3	1.26	0.94	0.55	1.53	1.29	1.14	1.13	1.40	1.20	0.85	0.50	0.42	0.70	0.89	1.04	0.70	0.76	0.94	1.64	0.59	0.45	0.72	0.52	1.46	1.20	0.61	0.88	0.93	0.81	1.46	1.81	1.11
156	p1H10	1.82	1.30	0.97	1.15	1.65	1.18	0.59	0.72	0.53	1.89	1.64	1.05	0.75	1.07	0.76	0.71	0.85	1.50	1.17	0.73	0.92	0.90	1.14	4.02	4.49	4.46	0.33	0.26	0.34	1.22	1.28	0.89
158	p1H18	0.52	0.65	1.16	0.83	0.67	0.23	1.49	1.33	1.35	1.34	1.21	0.97	1.51	2.07	0.70	0.67	0.22	1.07	0.49	0.15	0.65	0.80	1.00	1.71	1.19	0.75	1.47	0.95	0.79	1.78	1.45	1.00
160	p1H24	7.02	5.55	3.73	0.83	3.66	1.22	1.04	1.32	0.77	1.27	1.71	1.53	0.89	0.92	3.02	4.58	4.74	0.93	0.84	0.50	0.76	0.89	1.07	1.06	0.87	0.78	0.90	0.62	0.50	1.24	1.24	0.87
162	p1E22	0.48	0.30	0.53	1.93	2.28	2.13	1.68	1.65	2.38	0.41	0.35	0.49	0.35	0.46	1.26	1.00	1.62	1.04	1.05	0.62	0.85	1.66	1.16	1.07	1.02	0.90	1.01	1.27	1.27	1.05	1.39	0.89
164	p1H21	0.82	0.56	0.66	3.63	2.47	3.49	0.83	0.96	0.82	0.64	0.67	0.76	5.64	10.2	6.34	16.3	12.1	0.34	0.94	0.49	0.32	0.82	2.11	0.63	1.14	1.03	0.80	1.43	2.23	1.28	1.97	1.15
166	p1I1	1.64	0.90	1.61	1.02	1.71	0.85	1.13	1.33	0.97	0.88	0.85	0.77	0.78	0.84	1.43	2.04	1.37	0.52	0.45	0.30	1.22	1.28	1.73	0.98	0.70	0.51	0.66	0.51	0.77	2.16	2.43	1.44
168	p1H14	0.86	0.45	0.71	1.62	1.10	0.78	1.65	1.91	1.77	0.87	0.87	1.00	0.32	0.46	0.39	0.37	0.35	1.14	1.18	1.02	0.30	0.42	0.37	3.00	3.19	1.59	1.71	2.48	2.09	1.35	1.56	0.99
170	p1H11	0.96	0.63	0.72	2.19	2.53	2.46	2.53	3.11	3.04	0.74	0.67	0.79	0.46	0.87	1.11	0.78	0.74	0.24	0.41	0.22	0.42	0.68	0.92	1.28	1.28	1.26	2.20	2.56	1.81	1.12	1.13	0.97
172	p1H17	0.58	0.69	1.17	0.65	0.64	0.17	1.67	1.53	1.54	1.36	1.39	1.11	1.30	1.73	0.63	0.71	0.24	1.28	0.58	0.15	0.77	0.80	0.85	1.58	0.98	0.67	1.43	1.10	0.82	1.49	1.02	0.71
174	p1H12	2.06	1.50	1.29	1.79	2.33	2.36	0.84	1.09	0.87	0.80	0.63	0.69	0.82	2.34	0.96	0.78	1.03	1.02	1.56	0.71	0.58	0.95	0.61	0.93	1.28	1.16	0.51	0.63	0.76	2.85	3.33	2.57
176	p1H7	0.70	0.64	1.12	1.21	1.38	0.75	1.66	1.47	1.38	1.10	1.04	0.79	1.27	1.86	0.52	0.55	0.31	0.35	0.20	0.11	0.54	0.71	0.76	1.81	1.40	1.00	1.11	0.77	0.96	2.27	1.45	1.10

178	p1H15	2.34	1.79	1.89	0.91	1.77	0.73	0.92	0.85	0.57	1.17	0.80	0.76	1.10	1.14	1.43	2.36	2.35	0.61	0.32	0.18	1.05	1.30	1.88	1.00	0.73	0.43	0.83	0.48	0.37	1.77	1.51	0.96
180	p1H20	0.42	0.53	0.98	0.91	0.81	0.24	1.47	1.48	1.33	1.26	1.07	0.88	1.04	1.83	0.73	0.76	0.25	1.42	0.56	0.18	0.72	0.91	0.93	1.69	1.41	1.01	1.29	0.91	0.81	2.25	1.41	1.08
182	p1H8	0.86	0.87	1.87	1.02	1.13	0.22	2.17	2.37	1.65	2.56	2.34	1.63	1.93	2.30	1.00	0.97	0.42	0.15	0.06	0.03	0.94	0.98	1.07	0.87	0.57	0.32	0.78	0.44	0.67	2.28	1.78	1.17
184	p1H16	0.68	0.42	0.54	1.59	1.56	1.68	1.28	1.01	0.97	0.25	0.14	0.29	1.47	2.48	0.62	0.81	0.76	0.85	0.99	0.58	0.93	1.24	0.88	1.55	1.06	1.70	0.63	0.51	1.04	1.51	1.13	1.21
186	p1H9	1.31	0.64	0.55	0.94	0.74	0.46	1.88	2.01	1.52	1.05	0.97	1.02	0.25	0.40	0.45	0.35	0.47	1.26	1.82	1.00	0.38	0.46	0.73	3.02	1.86	2.13	2.67	3.21	2.86	1.66	1.28	0.87
188	p1H23	1.10	0.60	0.60	1.28	0.97	0.67	1.96	2.81	2.33	1.20	0.93	0.86	0.40	0.52	0.39	0.34	0.38	1.16	1.05	0.99	0.38	0.46	0.53	4.59	3.73	1.67	1.90	4.02	2.27	1.30	1.08	0.77
190	p1H10	7.36	4.40	4.02	1.57	7.80	2.18	0.85	1.11	0.73	0.97	0.82	0.98	1.07	1.42	2.62	3.20	3.68	0.65	0.65	0.43	0.69	0.81	1.24	0.73	0.80	0.71	0.35	0.39	0.50	1.91	1.81	2.21
192	p1H6	8.14	6.12	3.46	0.48	4.02	1.26	0.41	0.47	0.29	1.51	1.62	1.06	1.84	1.16	4.77	5.40	5.08	0.16	0.13	0.10	1.87	1.88	2.03	0.17	0.25	0.13	0.13	0.10	0.12	0.85	1.17	0.85
194	p1H13	0.59	2.10	1.03	0.83	0.58	0.21	2.87	1.67	1.71	0.49	0.64	1.22	3.85	5.85	1.99	1.55	0.72	4.92	3.05	0.37	2.06	1.40	1.91	0.51	0.46	0.90	0.68	0.37	0.49	1.05	0.47	0.87
196	p1H19	0.47	0.46	0.54	1.12	0.54	0.16	2.12	2.58	1.50	1.53	1.40	1.04	1.65	1.82	1.03	0.80	0.80	2.90	2.66	0.49	0.59	0.89	1.56	0.62	0.44	0.82	2.07	1.40	0.59	0.86	1.61	0.95
198	p1G22	1.14	0.78	0.72	2.75	2.67	1.41	8.35	12.7	8.79	1.35	1.21	1.01	0.59	0.70	0.77	0.63	0.77	0.58	0.60	0.24	0.33	0.47	0.49	0.94	0.72	1.03	1.65	1.71	1.62	1.88	2.51	2.06
200	p1G21	0.99	1.01	0.55	2.02	1.48	1.10	1.54	1.13	0.80	0.78	0.65	0.60	0.72	1.35	1.28	0.69	0.74	1.03	1.34	0.49	0.38	0.57	0.73	0.92	0.65	1.04	1.47	1.29	1.15	1.20	1.47	1.03
202	p1H1	1.33	0.82	0.82	2.50	2.70	1.59	0.94	1.26	1.95	1.22	1.06	0.98	0.81	0.87	0.65	0.82	1.08	0.46	0.38	0.25	0.65	0.94	1.02	1.44	1.84	1.44	1.77	1.69	1.29	1.01	0.97	0.81
204	p1G20	0.74	0.48	0.50	1.54	0.66	0.36	1.64	1.29	1.03	0.99	0.84	0.84	1.34	1.19	1.29	0.98	1.09	1.85	1.50	0.64	1.27	1.46	1.03	0.85	0.67	0.37	1.51	1.05	0.70	1.01	1.11	0.85
206	p1H5	2.46	1.48	1.52	1.40	3.58	1.14	5.93	16.5	11.3	0.97	0.81	0.87	0.80	1.19	0.92	1.18	1.14	0.49	0.58	0.40	0.69	0.78	0.91	1.07	0.84	1.04	0.55	0.44	0.29	2.09	2.45	2.06

208	p1G19	0.44	0.52	0.79	0.88	0.81	0.30	1.75	1.77	1.18	1.41	1.37	1.06	1.76	2.43	0.75	0.64	0.29	2.03	1.39	0.36	0.48	0.51	0.35	1.14	0.90	1.76	1.78	1.49	2.25	1.00	1.01	0.56
210	p1H2	1.13	0.55	0.49	0.84	0.83	0.39	2.34	4.04	4.03	0.68	0.56	0.47	0.12	0.17	2.43	2.35	1.86	0.91	0.82	0.72	1.94	2.47	2.65	2.55	2.29	0.35	1.66	2.25	1.06	0.61	0.57	0.36
212	p1G18	0.41	0.82	0.75	0.95	0.63	0.22	2.50	2.28	1.45	0.59	1.02	1.02	1.67	2.11	1.29	0.91	0.90	3.11	3.14	0.50	1.10	1.18	1.51	0.54	0.81	1.17	1.05	0.84	1.08	0.79	1.11	0.74
214	p1H4	1.61	1.00	1.44	0.92	1.89	0.63	1.15	1.46	1.00	1.26	1.15	1.02	0.90	0.66	1.11	1.21	1.26	0.21	0.08	0.04	1.06	1.04	1.37	0.81	0.78	0.44	0.69	0.54	0.45	1.26	0.89	0.74
216	p1H3	1.28	0.68	1.11	1.13	1.55	0.48	0.87	0.90	0.56	1.01	0.77	0.66	1.05	0.88	2.01	3.04	2.12	0.84	0.35	0.14	1.34	1.57	2.53	1.32	0.67	0.44	1.11	0.48	0.44	2.21	1.52	0.90
222	p1P3	1.32	0.79	1.45	1.39	2.10	1.47	17.1	10.7	10.8	0.82	1.02	1.15	0.53	0.83	0.48	1.08	1.92	0.44	0.79	0.38	0.27	0.44	0.66	0.47	0.55	0.30	7.15	7.93	6.97	1.12	1.18	1.12
224	p1A9	1.09	2.80	2.72	1.01	1.15	2.43	3.64	5.80	3.79	0.32	0.88	1.62	0.59	1.33	0.24	0.44	0.79	1.08	2.70	3.21	0.06	0.35	0.47	0.16	0.20	1.42	0.79	1.64	3.82	0.72	1.78	3.33
224	p1A8	0.52	1.10	1.43	0.60	0.57	0.83	4.46	4.51	2.81	0.17	0.42	1.28	1.75	4.84	0.99	2.32	6.46	4.52	37.2	24.2	0.63	1.17	1.40	0.13	0.14	0.44	0.27	0.49	1.37	0.39	0.56	1.59
226	p1B17	1.04	24.7	2.25	1.16	1.22	1.25	1.08	0.99	0.81	0.69	0.81	0.79	1.50	2.35	2.53	2.56	6.59	27.2	223	109	0.51	1.00	0.68	0.81	0.85	0.86	0.83	1.22	2.43	0.86	1.55	1.09
228	p1O20	0.58	1.42	5.10	0.73	4.10	7.70	0.63	0.84	1.02	0.74	2.32	2.75	0.35	0.96	0.06	0.10	0.27	4.25	7.39	5.74	0.09	0.44	0.44	1.19	14.2	31.1	5.43	8.1	15.1	0.68	2.46	5.18
230	p1B2	0.66	1.96	5.58	0.58	1.02	1.50	1.53	0.95	0.87	0.57	2.55	4.40	0.56	1.20	0.25	1.03	0.99	7.01	22.2	28.8	0.33	0.98	1.70	0.44	0.54	2.32	0.67	4.52	13.0	0.51	1.10	1.44
232	p1B3	1.24	1.35	3.82	1.78	2.17	4.64	3.30	3.42	2.96	0.79	1.01	1.32	0.10	0.26	0.15	0.40	0.85	0.33	1.27	1.04	0.09	0.42	1.08	0.33	0.41	3.73	0.97	1.72	1.54	0.84	1.63	2.88
236	p1B19	1.34	1.54	5.26	1.28	0.89	2.48	3.63	3.26	6.08	0.66	2.09	5.36	0.72	1.85	0.01	0.03	0.07	0.36	2.23	1.77	0.01	0.03	0.06	0.01	0.01	0.01	1.23	0.69	3.23	0.80	2.03	1.54
236	p1B18	0.65	18.3	2.73	1.68	1.06	2.54	26.2	10.0	4.27	0.19	0.57	5.12	2.30	16.1	0.16	0.34	0.89	2.35	29.2	13.7	0.20	0.34	0.39	0.17	0.17	0.15	0.21	0.17	0.69	1.36	0.66	1.75
238	p1N17	2.81	1.63	2.01	1.24	3.19	1.83	4.80	0.87	0.57	1.48	1.29	1.16	0.47	0.59	1.06	1.27	1.67	3.14	3.81	8.23	0.41	0.43	0.63	0.35	0.37	0.23	0.16	0.18	0.21	0.71	0.86	0.73

240	p1A24	4.49	3.48	2.62	0.34	1.91	1.06	0.94	0.99	0.59	1.36	1.63	1.94	6.97	14.2	1.24	2.41	3.26	0.37	1.01	0.74	0.44	0.77	1.17	0.10	0.13	0.08	1.85	0.91	2.15	0.66	1.03	0.82
244	p1B1	1.45	1.74	2.07	0.14	0.16	0.32	1.08	1.32	0.86	1.03	1.09	1.69	6.95	24.7	0.04	1.06	1.42	0.65	2.03	1.31	0.01	0.74	1.05	0.03	0.02	0.04	2.53	1.72	2.76	0.61	1.04	0.88
248	p1A4	0.63	0.55	0.78	2.47	3.35	4.34	2.21	1.60	3.18	0.92	0.59	1.17	0.30	0.55	0.35	0.33	0.84	0.22	1.78	1.26	0.31	0.89	0.94	1.14	1.97	0.82	2.22	2.17	3.25	1.70	2.50	2.14
248	p1A2	0.57	0.81	1.07	1.36	1.62	1.87	1.48	1.73	1.65	0.44	0.48	1.02	0.33	0.65	0.51	1.49	3.47	0.51	1.21	0.87	0.33	2.30	2.64	0.86	1.50	0.70	0.91	1.32	2.58	0.58	0.80	0.86
248	p1A3	0.55	1.18	0.80	0.56	1.04	1.29	1.42	1.47	1.06	0.56	0.92	1.40	0.20	0.36	0.37	1.88	5.15	1.31	31.9	18.8	0.41	4.16	5.14	1.20	1.46	0.63	0.63	1.50	2.78	0.52	0.97	1.02
248	p1A1	0.58	1.50	0.66	1.06	1.04	1.19	1.35	1.20	1.07	0.77	1.65	0.82	0.59	1.31	0.53	1.76	10.0	1.02	49.3	22.8	0.76	4.24	4.31	0.75	0.62	0.72	0.53	0.70	2.19	0.87	0.94	1.07
250	p1A16	0.75	1.16	0.95	1.36	1.29	1.26	1.22	1.65	1.37	0.66	0.89	0.74	0.63	1.20	0.44	0.92	5.53	0.80	9.87	5.89	0.39	1.20	1.04	0.55	0.56	1.34	0.58	0.85	0.22	0.91	1.02	1.08
250	p1A15	0.49	1.03	1.93	0.27	1.07	1.07	0.67	1.54	1.03	0.39	1.66	1.22	0.04	0.10	0.79	3.65	6.09	0.90	5.67	5.24	0.29	2.68	3.57	0.83	2.06	13.2	0.73	2.53	7.79	0.18	1.20	1.04
250	p1A18	0.60	0.86	1.64	0.28	1.02	0.96	0.70	1.44	1.22	0.46	1.52	1.35	0.23	0.26	0.96	4.48	4.95	1.01	4.25	7.77	0.50	2.95	4.85	2.05	8.73	12.21	1.02	2.75	7.82	0.31	0.84	0.85
250	p1A17	0.78	2.06	0.51	1.32	1.13	0.93	0.79	1.03	1.54	0.64	1.17	0.74	0.47	0.67	1.01	6.98	22.0	1.16	34.1	13.81	1.85	4.66	2.89	1.18	0.95	0.90	0.76	1.10	2.00	1.02	1.40	1.08
252	p1B14	0.78	1.20	5.92	6.75	3.77	5.37	1.13	5.81	5.57	0.17	0.33	0.66	2.16	4.85	0.10	0.12	0.22	0.45	0.65	1.54	0.84	5.03	2.44	0.04	0.05	0.05	2.81	1.35	4.13	0.30	0.29	0.30
252	p1B15	1.22	1.15	7.37	14.4	11.6	11.0	0.92	6.90	6.45	0.40	0.61	0.94	1.89	3.76	0.08	0.10	0.24	0.34	0.46	0.83	0.54	4.88	2.25	0.05	0.05	0.06	8.90	3.01	8.69	1.01	0.68	0.91
252	p1B16	1.14	0.97	7.33	13.6	10.7	9.97	1.04	7.34	7.64	0.45	0.65	1.03	1.99	3.48	0.09	0.10	0.24	0.47	0.45	1.35	0.78	5.19	2.01	0.05	0.04	0.06	8.93	3.67	13.9	0.95	0.67	0.82
254	p1A12	0.44	0.91	1.17	1.07	0.92	1.33	4.01	3.77	3.56	0.77	1.59	2.04	0.37	0.55	0.44	0.70	1.06	2.50	3.37	4.51	0.19	0.38	1.19	0.37	0.63	2.83	0.92	2.17	3.37	0.60	1.03	1.69
254	p1A11	0.16	3.67	0.42	0.58	0.42	0.57	2.87	1.45	0.95	0.20	0.43	2.01	1.85	5.50	1.17	5.82	26.9	6.17	142	62.4	2.33	5.10	6.21	0.13	0.22	0.81	0.17	0.44	1.01	0.53	0.29	1.22

256	p1A13	0.34	2.14	1.10	0.42	0.36	0.65	3.57	3.81	3.15	0.19	0.79	1.98	0.34	0.80	1.24	1.86	3.13	1.49	4.12	4.95	0.92	1.18	2.38	0.06	0.06	1.27	0.29	0.81	2.00	0.21	0.36	1.05
258	p1A14	0.27	5.92	0.30	0.40	0.45	0.71	7.23	2.92	2.30	0.08	0.18	1.06	2.39	6.30	4.94	7.07	7.75	6.29	28.0	29.3	4.13	5.41	7.32	0.17	0.24	0.33	0.11	0.23	0.77	0.59	0.39	0.89
260	p1A19	0.56	2.88	0.45	0.99	0.90	1.33	2.11	1.14	0.73	0.45	0.53	0.82	1.46	4.56	4.90	1.37	15.7	1.21	33.7	4.98	0.74	5.11	2.30	0.68	0.81	0.61	0.43	0.56	0.52	0.86	0.64	0.89
262	p1A20	0.20	8.18	0.34	0.36	0.40	0.68	14.1	8.24	6.34	0.06	0.27	1.66	1.38	3.74	2.34	3.16	4.93	3.33	10.0	7.74	0.98	1.70	2.70	0.08	0.08	0.12	0.08	0.14	0.57	0.38	0.18	1.26
264	p1A22	0.78	1.44	2.34	0.52	1.40	1.52	2.59	3.81	2.38	0.46	1.47	1.92	0.77	0.98	0.19	0.56	0.87	0.84	1.41	1.18	0.07	0.24	0.32	0.61	1.92	3.26	2.03	4.60	4.51	0.75	2.15	3.17
266	p1A23	0.71	2.10	1.41	0.07	0.06	0.13	3.12	2.80	1.63	1.06	1.94	2.95	7.39	18.1	0.03	0.95	1.14	1.34	4.67	2.78	0.02	0.79	1.41	0.01	0.01	0.01	1.17	0.96	2.10	0.13	0.23	0.21
268	p1B21	3.08	1.82	2.69	0.59	3.22	0.71	0.56	0.66	0.41	1.48	1.45	0.75	0.77	0.99	96.6	117	165	0.17	0.31	0.17	1.35	3.24	55.7	0.37	0.40	0.27	28.9	30.2	86.8	0.90	1.00	0.67
268	p1B20	0.72	0.85	0.77	1.17	1.27	1.05	0.74	1.03	0.57	0.80	0.77	0.64	0.61	1.25	1505	1742	2952	1.05	3.07	1.20	17.5	32.8	682	0.76	0.85	1.29	45.1	42.8	224	1.14	0.96	1.15
270	p1C17	0.38	0.91	0.47	1.39	0.63	1.21	3.27	2.63	2.74	0.37	0.28	0.59	0.21	0.18	3.94	4.47	5.82	3.02	5.66	2.13	4.54	4.06	4.88	0.46	0.55	0.59	0.80	0.95	1.70	0.92	0.83	1.03
270	p1C18	0.42	0.76	0.51	1.51	0.99	1.69	2.88	2.45	2.64	0.32	0.31	0.54	0.16	0.27	4.76	5.23	8.06	2.62	7.14	2.17	3.71	3.18	4.70	0.45	0.66	0.69	0.67	1.03	1.59	0.91	1.00	0.96
272	p1D8	2.74	2.48	2.75	0.99	2.72	1.61	0.74	0.78	0.74	1.31	1.08	1.01	0.84	0.91	2.08	2.03	2.31	0.09	0.17	0.18	1.11	1.09	1.50	0.62	0.97	0.73	0.27	0.44	0.80	1.16	1.31	1.07
274	p1A10	0.13	1.74	1.78	0.33	2.02	1.93	3.00	3.12	2.80	0.20	1.69	2.86	0.06	0.27	0.30	1.81	3.43	0.14	2.68	4.91	0.12	0.78	1.90	0.53	0.36	6.58	0.90	5.24	22.3	0.10	0.91	2.13
276	p1G24	1.41	1.31	1.91	3.09	4.24	2.63	0.92	1.60	0.93	1.14	1.39	1.75	0.14	0.30	0.40	0.70	1.02	0.42	1.73	0.82	0.17	0.45	0.48	0.62	1.04	1.26	0.48	1.02	1.36	2.79	4.68	4.49
278	p1G23	2.81	1.41	1.97	1.06	2.80	1.23	0.90	1.14	0.92	0.94	0.85	0.80	0.37	0.51	2.52	1.37	2.13	0.90	2.33	0.80	0.73	1.13	1.75	0.66	0.90	0.81	0.44	0.84	0.95	1.72	2.05	1.83
280	p1G5	0.34	0.86	1.46	0.56	1.70	1.61	0.63	1.16	1.35	0.43	1.23	1.77	0.21	0.53	0.21	0.97	0.90	1.21	3.69	3.47	0.45	1.28	2.11	0.89	2.10	4.01	0.68	2.58	5.12	0.36	1.31	1.57

282	p1G7	0.74	0.61	0.80	1.25	1.26	1.03	1.13	1.77	1.39	0.87	1.28	1.54	0.37	0.56	0.36	0.61	0.52	1.12	1.85	1.19	0.23	0.40	0.50	1.27	0.92	1.35	1.62	2.39	1.90	0.94	1.54	1.32
284	p2A23	1.79	1.24	1.15	1.81	2.26	1.79	1.11	1.28	0.95	1.26	0.87	0.77	0.77	1.43	0.76	0.91	2.36	0.59	0.61	0.40	0.32	0.40	0.35	1.32	1.14	0.79	0.85	0.76	0.66	1.68	2.18	1.53
286	p1G1	0.54	0.39	0.81	1.68	2.29	3.15	1.05	1.41	1.33	1.85	2.65	1.71	0.35	0.58	0.21	0.33	0.35	0.53	0.85	0.75	1.04	1.89	1.39	1.17	1.16	0.98	0.74	0.91	0.95	1.16	1.20	1.52
288	p1G15	7.04	5.26	3.09	0.70	3.29	2.10	0.57	0.77	0.55	2.10	1.82	1.03	1.54	1.44	3.27	3.54	4.87	0.34	0.56	0.36	1.92	1.93	1.93	0.33	0.63	0.51	0.20	0.28	0.32	0.80	1.05	0.89
290	p1F23	0.90	0.42	1.09	1.89	1.13	1.09	1.22	2.57	1.79	1.10	0.88	0.70	0.42	0.55	0.33	0.40	0.32	0.84	0.77	0.70	0.29	0.32	0.32	1.50	1.79	1.29	2.59	2.59	1.90	1.67	1.38	1.16
292	p1G8	1.02	0.70	1.33	0.86	3.03	1.31	1.15	3.36	2.20	1.04	1.93	1.56	0.33	0.71	0.23	0.35	0.43	0.49	0.79	0.86	0.20	0.62	0.52	13.4	14.8	10.4	1.32	4.10	3.54	0.96	2.78	2.43
294	p1G13	8.69	6.51	4.23	1.06	4.37	2.12	0.37	0.47	0.36	1.39	1.41	0.82	1.03	1.17	5.78	7.17	9.42	0.46	0.47	0.43	2.39	2.87	3.25	0.15	0.16	0.13	0.28	0.43	0.64	0.79	1.19	0.96
296	p1G10	1.01	0.61	1.52	1.69	1.76	1.65	2.01	3.58	2.49	1.32	1.13	0.99	0.44	0.65	0.44	0.62	0.49	0.67	0.68	0.40	0.30	0.44	0.37	1.46	0.96	0.72	2.48	2.48	1.88	1.56	1.51	1.21
298	p1F24	1.17	0.70	1.01	9.63	5.52	8.15	0.56	0.87	0.56	12.8	11.5	14.2	0.12	0.15	0.11	0.14	0.08	2.07	1.62	1.89	0.13	0.17	0.19	0.45	0.58	0.36	1.18	1.06	1.55	5.29	5.20	4.26
300	p1G2	1.69	1.00	1.36	5.01	4.65	3.35	1.04	1.47	0.83	1.28	1.39	1.24	0.36	1.37	0.42	0.48	0.82	0.61	0.84	0.43	0.28	0.29	0.48	0.89	0.65	1.21	0.91	0.92	0.69	2.10	2.84	2.62
302	p1G11	0.53	0.42	0.42	0.55	0.86	0.65	1.43	3.83	3.56	0.45	0.36	0.43	1.21	1.55	0.29	0.53	0.65	3.60	4.27	4.45	0.16	0.26	0.20	5.96	8.65	6.54	11.3	18.7	26.3	0.98	1.30	1.18
304	p1G16	0.53	0.29	0.55	1.58	1.64	1.65	1.10	2.51	1.64	0.87	0.64	0.71	0.41	0.74	0.23	0.34	0.36	0.71	1.07	0.87	0.87	1.07	1.38	8.25	8.61	5.74	0.97	2.34	3.68	1.07	1.21	1.14
306	p1G9	1.32	0.61	0.77	1.07	1.78	0.79	0.74	1.22	0.96	0.95	0.85	0.66	0.79	1.09	0.94	1.06	1.43	1.28	1.06	1.00	0.49	1.14	1.76	0.85	0.66	0.52	1.39	1.80	2.13	0.98	1.04	0.96
308	p1G4	0.83	0.42	0.92	1.67	3.65	3.50	0.69	0.62	0.54	1.48	0.87	1.17	0.49	0.74	0.71	1.39	2.65	0.95	1.70	1.60	0.56	0.77	0.78	2.02	0.90	0.45	1.44	3.73	5.27	4.99	7.58	9.71
310	p1G14	3.40	5.55	4.36	0.70	4.34	1.75	0.47	0.56	0.36	1.14	0.98	0.90	1.30	1.11	3.49	4.03	4.68	0.15	0.16	0.13	1.08	1.48	1.63	0.33	0.49	0.25	0.17	0.18	0.18	0.94	1.57	1.09

	p1A6	0.30	0.59	2.57	0.31	1.10	1.88	1.02	1.31	1.44	0.56	1.50	2.81	0.16	0.24	0.23	0.80	1.29	0.55	2.17	2.92	0.15	0.91	1.99	0.99	2.21	1.94	1.18	3.56	5.14	0.40	1.08	1.87
	p1A5	5.61	2.87	4.93	1.15	4.00	2.95	0.60	0.87	0.47	1.77	1.64	1.79	0.56	0.55	1.70	3.28	3.93	0.19	0.23	0.22	0.98	1.78	3.76	0.25	0.33	0.24	0.15	0.20	0.26	1.25	1.65	1.39
	p1B9	1.14	19.9	3.20	0.86	2.05	1.01	1.01	1.35	1.08	0.63	3.53	5.52	0.75	2.19	1.86	2.87	4.45	0.17	0.47	0.17	5.52	18.9	19.0	0.16	0.16	0.41	0.29	0.57	4.92	0.32	0.61	1.23
	p1B6	0.69	18.7	1.14	0.59	0.72	0.92	0.70	0.54	0.61	0.59	0.89	1.01	1.54	4.45	2.13	3.09	8.12	1.22	2.38	1.81	10.0	25.1	16.0	0.60	0.57	0.58	0.99	1.38	2.38	0.79	0.54	0.87
	p1B8	1.11	35.1	0.81	0.74	1.23	1.02	1.15	1.35	0.91	0.57	1.27	3.41	0.85	3.57	2.53	3.71	7.34	0.64	1.60	1.29	5.86	21.3	21.4	0.55	0.47	0.46	0.39	0.44	1.19	0.43	0.55	1.38
	p1B7	0.60	34.9	1.17	0.59	0.74	0.80	0.90	0.80	0.72	0.48	0.84	1.76	1.94	4.62	4.60	7.31	12.00	1.15	2.78	1.89	21.1	41.1	42.4	0.46	0.44	0.57	0.75	0.85	1.97	0.46	0.40	0.97
	p1G17	0.85	1.53	0.92	1.10	0.83	1.05	2.04	1.86	1.78	1.40	2.19	2.22	0.81	0.87	0.46	0.74	0.93	1.06	1.46	0.76	0.24	0.44	0.99	1.10	0.77	0.82	2.00	2.84	3.17	0.96	0.97	1.00
	p1G3	0.38	0.39	1.02	0.76	1.16	0.87	1.38	1.89	1.71	0.81	0.73	1.00	0.81	2.03	0.24	0.48	0.30	1.65	1.68	1.50	0.58	0.86	1.02	0.95	0.99	1.52	1.23	1.74	3.24	1.28	0.97	1.78
	p1F22	0.94	0.56	0.75	1.15	1.24	1.60	2.56	3.11	2.46	1.10	1.70	1.67	0.45	0.47	0.29	0.31	0.35	0.74	0.85	0.83	0.16	0.27	0.31	1.51	1.59	0.85	2.54	3.06	3.72	1.03	1.01	0.75
	p1G12	0.29	0.41	0.96	0.99	1.95	1.56	0.80	1.75	1.22	0.54	0.87	1.24	0.12	0.31	0.16	0.30	0.39	1.53	2.64	2.85	0.33	0.51	0.84	1.69	3.85	4.80	0.83	3.71	5.25	1.74	2.60	3.23
	p1F11	1.49	0.94	1.01	1.16	1.64	1.56	0.64	0.91	0.71	0.84	0.69	1.00	0.77	1.16	0.53	1.01	0.89	0.64	0.85	0.57	0.98	1.77	1.64	2.48	2.83	1.47	0.46	0.74	0.73	1.62	1.80	1.81
	p1F16	2.63	0.62	0.65	6.26	1.77	1.43	0.39	0.16	0.13	1.07	0.39	0.30	0.29	0.30	2.73	3.22	3.38	1.02	0.52	0.21	7.19	6.39	9.11	0.31	0.33	0.16	1.44	0.39	0.29	1.68	1.44	1.20
	p1F14	0.75	0.39	0.89	0.97	0.92	0.84	1.29	1.35	1.36	0.93	0.90	1.03	0.54	0.98	0.16	0.22	0.30	1.52	3.98	3.61	1.00	1.72	1.66	0.06	0.07	0.07	3.27	4.66	10.6	1.25	1.05	1.46
	p1F17	1.52	0.95	1.82	5.53	3.68	5.49	0.24	0.32	0.33	2.14	2.44	2.18	0.86	1.66	0.22	0.92	1.32	11.7	16.0	9.79	0.04	0.27	1.01	0.06	0.07	0.07	0.12	0.12	0.23	1.19	2.33	1.66
	p1C2	0.26	0.31	0.31	1.00	1.28	1.53	5.09	13.6	16.1	0.17	0.21	0.31	0.19	0.37	0.67	1.53	2.00	0.60	2.84	1.53	7.05	9.91	9.48	0.95	18.0	39.5	0.27	1.57	3.32	0.52	0.80	0.64

334	p1F3	1.64	1.06	1.20	1.52	1.52	1.41	1.05	1.55	1.28	1.26	1.02	0.99	0.71	1.05	0.46	0.50	0.87	0.56	0.75	0.45	0.40	0.64	0.66	1.14	0.99	0.74	0.79	1.38	1.14	1.82	2.15	1.65
336	p1F20	0.69	1.35	1.32	1.10	1.16	1.03	1.73	2.66	2.42	0.34	0.64	0.76	0.86	1.82	0.53	0.87	0.98	2.60	4.08	2.62	1.11	1.52	2.07	0.27	0.27	0.71	0.37	0.71	1.14	0.70	0.89	1.08
338	p1F6	1.92	1.67	2.74	0.11	0.17	0.28	1.73	1.30	1.02	1.80	2.33	2.74	5.33	7.33	0.04	0.73	0.91	0.45	1.45	0.91	0.01	0.45	0.61	0.04	0.05	0.04	2.18	2.32	3.68	0.92	1.03	0.84
340	p1F4	1.50	1.06	0.87	1.14	1.59	1.15	0.68	1.05	0.83	0.92	0.72	0.66	0.94	1.85	2.78	6.20	5.14	2.83	2.10	1.95	0.44	0.65	0.99	0.87	0.89	0.78	0.73	0.62	0.82	1.34	1.50	1.15
342	p1F15	1.29	0.82	1.10	1.97	1.87	1.33	1.71	1.75	1.30	0.62	0.78	0.74	0.34	0.81	0.63	0.83	1.06	0.79	1.22	0.71	0.35	0.56	0.84	1.57	1.93	0.90	1.16	1.46	2.33	1.36	1.38	0.93
344	p1F13	0.32	1.18	2.65	0.54	3.15	2.25	0.74	1.19	1.19	0.80	3.41	4.43	0.21	0.49	0.17	0.52	0.71	0.43	1.58	1.80	0.28	1.94	2.63	2.51	14.4	12.9	0.82	6.59	7.42	1.07	3.89	3.99
346	p1A7	0.74	0.60	0.50	0.87	0.63	0.80	0.85	0.95	0.69	0.71	0.57	0.76	0.73	1.22	1.59	1.11	2.67	5.16	9.61	4.32	1.15	2.12	2.18	0.99	0.37	0.61	1.58	1.30	1.60	1.86	2.36	1.59
348	p1A21	1.35	0.89	1.29	1.27	1.17	0.98	0.85	1.49	1.32	1.07	1.31	1.17	0.52	0.49	0.18	0.22	0.24	0.86	1.13	1.08	0.18	0.26	0.36	0.61	0.33	0.38	1.19	1.97	1.36	2.27	2.95	3.16
350	p1B5	1.08	0.94	3.22	1.73	1.39	3.45	2.08	2.24	1.78	0.34	0.55	1.57	0.10	0.17	0.07	0.13	0.10	0.90	1.37	1.04	0.02	0.03	0.06	0.18	0.30	1.85	1.47	1.82	3.02	1.66	3.19	3.40
350	p1B4	0.71	7.10	2.98	2.43	1.47	3.12	10.3	7.97	5.70	0.48	0.93	2.72	0.65	1.12	0.44	0.75	0.87	3.69	9.15	4.61	0.25	0.22	0.48	0.32	0.30	0.80	0.52	0.79	2.38	1.93	1.75	5.01
352	p1B12	1.74	1.15	1.23	0.70	1.53	0.93	0.60	0.57	0.30	0.94	0.88	0.89	0.34	0.37	1.15	1.50	1.82	0.74	0.84	1.09	0.75	1.02	0.97	0.49	0.64	0.37	1.28	1.90	1.56	1.10	1.37	1.41
352	p1B11	0.21	0.17	0.40	1.25	1.64	1.18	1.14	0.66	0.45	0.94	1.02	1.31	0.28	0.30	0.27	0.64	1.45	1.42	2.25	1.85	0.06	0.10	0.22	0.93	1.11	1.39	2.23	3.90	3.56	2.04	2.69	3.58
352	p1B10	0.46	0.97	0.71	1.31	0.94	1.06	1.92	2.27	1.82	0.51	0.43	0.63	0.86	2.71	0.44	0.99	3.03	1.34	5.67	4.03	1.05	1.41	1.44	0.72	0.82	0.98	0.55	0.67	0.88	0.90	1.41	1.05
354	p1B13	0.93	0.99	1.06	0.85	1.01	0.91	0.77	0.97	0.68	1.00	1.03	1.02	1.22	1.61	1.20	1.42	1.66	1.24	1.26	0.83	1.96	1.86	2.20	0.55	0.76	0.41	0.31	0.29	0.39	0.98	0.94	1.04

356	p1B22	3.21	1.60	3.89	1.04	0.81	0.97	3.14	4.98	4.06	1.67	1.55	1.15	0.20	0.21	0.30	0.34	0.29	2.78	2.80	2.50	0.05	0.06	0.07	0.13	0.11	0.04	0.28	0.87	0.39	3.93	3.62	4.04
358	p1B23	0.36	0.52	1.08	0.80	1.09	1.17	0.93	0.66	0.85	0.27	0.52	0.82	1.05	1.03	17.2	12.5	8.51	1.84	6.94	6.08	1.79	2.46	3.57	0.49	0.59	0.97	0.52	1.43	4.14	0.58	0.86	0.83
360	p1B24	0.54	0.24	0.51	0.67	0.52	1.09	1.88	2.32	1.98	0.65	0.54	0.70	0.34	0.34	0.86	1.59	1.27	1.32	1.41	1.33	0.92	1.14	1.14	4.95	2.46	1.92	2.93	3.67	5.83	0.80	0.98	0.76
362	p1C3	0.71	0.68	0.77	2.10	0.92	1.22	4.25	2.15	1.97	1.07	1.18	1.09	0.46	0.54	0.48	0.57	0.55	1.86	2.73	1.77	0.73	1.09	1.29	0.18	0.10	0.36	1.15	1.00	1.09	1.01	2.33	1.40
364	p1C4	0.80	0.81	0.75	1.82	1.94	1.74	1.18	1.08	1.02	0.88	0.77	0.78	0.73	1.04	0.79	1.02	1.78	0.92	6.16	1.47	0.55	0.81	0.71	1.06	0.88	0.96	0.67	0.66	0.78	1.44	2.81	1.64
366	p1C5	0.52	0.67	1.21	0.93	0.89	1.07	1.80	2.92	2.03	0.70	1.42	1.68	0.58	1.01	0.61	0.85	0.89	0.72	1.39	1.15	0.39	0.96	1.04	1.46	1.37	1.15	1.58	3.45	4.04	0.73	1.21	0.77
368	p1C6	0.41	3.01	0.46	0.86	0.85	0.76	1.67	1.69	1.30	0.51	0.72	0.81	0.97	2.40	1.15	4.75	19.92	1.79	28.4	14.2	2.36	5.29	5.01	0.57	0.94	1.90	0.38	0.60	1.00	0.57	0.66	0.96
370	p1C7	0.93	2.60	0.63	0.74	0.89	0.75	3.35	1.93	1.49	0.61	0.93	0.83	1.81	2.86	1.92	1.25	1.69	2.98	11.76	0.31	1.90	1.61	1.45	0.68	0.64	0.58	0.51	0.85	0.76	0.84	1.00	0.74
372	p1C8	0.11	0.51	0.83	1.17	1.79	2.54	1.93	2.11	1.46	0.07	0.82	0.84	0.61	1.54	0.89	0.77	1.04	1.70	7.73	8.58	0.20	0.66	0.98	0.04	0.54	3.98	3.29	9.13	21.3	0.23	1.77	1.66
374	p1C9	1.04	1.00	1.67	1.31	1.62	1.61	1.18	2.32	1.60	0.75	0.86	0.89	0.42	0.57	0.57	0.60	1.00	0.81	1.62	1.21	0.84	0.80	1.01	0.46	0.38	0.49	0.92	1.27	1.68	1.15	1.33	1.56
376	p1C10	0.65	0.58	0.43	1.29	1.10	1.42	0.93	1.48	1.17	1.07	0.93	0.84	0.68	0.84	14.6	15.2	28.7	0.64	1.66	0.81	1.18	2.19	2.92	0.89	1.18	1.55	0.76	0.77	1.09	0.89	0.94	0.70
378	p1C11	0.69	1.30	0.86	1.59	0.56	0.90	2.64	2.63	2.42	0.45	0.52	1.32	0.79	1.37	0.90	0.59	1.45	1.73	11.0	5.75	0.58	0.87	0.78	0.13	0.07	0.56	0.81	1.46	3.60	1.00	1.63	1.11
380	p1C12	1.00	1.39	0.78	1.31	1.07	1.31	2.18	1.86	1.29	0.75	0.91	1.05	0.97	2.57	0.55	0.88	3.79	0.76	24.4	7.12	0.84	1.64	0.79	0.86	0.91	0.84	0.76	0.97	1.41	1.05	0.81	1.04
382	p1C13	0.61	5.60	1.61	0.78	0.80	0.90	3.98	3.29	2.33	0.73	2.76	5.00	0.54	1.10	1.24	1.66	2.43	0.98	2.13	1.13	1.32	2.88	3.71	0.33	0.41	0.60	0.37	0.55	0.94	1.14	0.95	2.34

384	p1C14	0.44	0.47	0.52	0.64	0.96	0.89	4.73	3.06	1.92	0.59	0.69	1.91	1.10	2.27	1.64	2.85	2.68	2.62	7.82	2.55	1.95	1.59	1.62	0.47	0.44	0.39	0.35	0.45	1.13	0.50	0.76	0.79
386	p1C15	1.40	0.82	0.84	0.50	0.65	0.50	2.02	5.84	2.20	2.00	1.39	1.64	0.37	0.42	0.46	0.58	0.66	0.61	0.97	0.72	0.47	0.73	0.78	30.7	28.4	10.3	2.22	5.85	5.44	1.82	1.79	1.18
388	p1C16	2.84	1.76	2.30	0.95	1.80	1.11	0.95	5.76	5.09	1.57	1.27	0.98	0.66	0.68	1.25	1.79	1.76	0.94	25.0	14.8	0.37	0.41	0.59	0.42	0.61	0.33	0.35	0.67	1.30	0.63	1.28	0.97
390	p1C19	2.33	1.63	1.78	0.58	0.90	0.67	0.73	1.35	1.04	0.98	0.97	0.98	0.33	0.33	13.7	17.4	26.3	4.16	5.72	3.15	1.55	1.72	4.66	0.72	0.74	0.80	0.76	1.04	1.09	0.50	0.61	0.65
392	p1C20	0.68	1.08	1.96	1.56	1.12	1.04	2.89	4.45	4.18	0.71	1.25	1.76	0.31	0.47	0.27	0.48	0.55	0.81	1.58	1.69	0.51	0.82	1.21	0.51	0.37	1.45	1.10	3.34	4.67	0.67	2.21	1.87
396	p1P5	3.26	1.18	1.71	70.5	33.8	21.0	1.93	6.73	7.47	1.84	2.20	1.55	0.75	0.67	1.12	0.57	0.53	0.17	0.23	0.12	0.94	0.49	0.27	0.53	0.47	0.33	6.09	2.15	4.70	0.79	0.78	0.59
398	p2L23	0.99	0.66	0.57	63.1	35.3	56.2	245	289	339	0.47	0.50	0.42	0.47	1.04	0.30	0.32	0.29	2.63	1.23	1.15	0.22	0.31	0.34	0.77	0.70	0.55	28.3	15.9	28.5	2.59	2.39	2.30
402	p1K9	1.09	0.61	1.08	1.23	0.55	1.15	1.42	2.46	2.34	2.23	1.48	1.83	0.01	0.01	0.18	0.24	0.25	1.96	1.86	1.27	0.26	0.44	0.57	0.01	0.00	0.00	0.82	1.72	0.90	1.97	2.33	1.62
404	p1K23	1.91	1.08	1.43	0.91	0.60	0.75	1.67	1.98	1.51	2.01	1.52	1.89	0.72	0.72	0.18	0.18	0.10	3.36	1.73	1.21	0.09	0.11	0.13	0.88	0.47	0.13	2.17	2.19	3.26	1.01	0.95	0.92
406	p1K15	2.31	1.24	1.63	1.23	2.38	1.19	0.89	1.09	0.82	1.11	0.96	0.84	2.57	2.98	13.2	14.6	17.8	0.43	0.63	0.43	0.68	1.01	4.72	0.43	0.50	0.26	0.28	0.31	0.36	1.25	1.49	1.08
408	p1K8	1.32	0.86	0.76	0.97	1.35	1.29	1.08	1.98	1.56	0.98	0.71	0.72	0.94	1.46	3.51	1.11	0.61	0.32	0.73	0.26	3.47	2.77	0.84	0.56	0.76	0.60	0.59	0.97	0.81	1.32	1.83	1.20
410	p1M24	1.71	0.94	1.37	1.88	2.31	1.49	1.24	1.49	1.29	0.81	0.74	1.10	0.42	0.75	1.22	0.69	1.28	0.71	1.79	0.89	0.55	1.55	1.03	0.78	0.83	0.70	0.47	0.89	0.90	1.71	3.29	1.88
412	p1K7	2.24	0.79	0.56	1.84	0.69	0.58	2.40	4.45	2.78	2.16	1.57	1.05	0.62	0.51	0.46	0.21	0.13	1.44	0.62	0.43	0.32	0.37	0.31	2.24	1.03	0.61	3.53	2.35	1.35	1.53	1.43	0.63
414	p1K16	1.23	0.58	0.37	1.14	0.64	0.49	2.19	2.31	1.76	1.05	0.69	0.64	0.48	0.50	0.41	0.19	0.15	0.96	1.55	1.43	0.20	0.16	0.11	3.56	2.37	1.22	2.65	3.65	1.49	1.13	1.07	0.48
416	p1K18	3.59	1.63	1.92	1.02	2.11	1.23	1.07	1.33	0.92	2.28	1.71	1.51	0.89	0.76	1.94	1.43	1.92	0.11	0.10	0.12	0.57	1.20	1.35	0.58	0.72	0.43	0.54	0.64	0.61	0.91	1.02	0.91
418	p1N1	0.83	0.42	0.54	1.45	0.67	0.85	1.38	1.52	1.32	1.28	0.96	1.03	0.24	0.36	0.24	0.13	0.16	0.86	0.97	0.88	0.14	0.28	0.30	2.03	1.33	1.80	1.75	2.03	1.64	1.84	1.86	1.24

420	p1K22	1.82	1.13	0.74	0.90	0.73	0.51	2.20	2.31	1.76	1.27	0.91	1.03	0.55	0.61	0.37	0.26	0.24	1.30	1.51	0.67	0.20	0.23	0.21	1.45	1.17	0.40	1.52	1.16	1.00	1.06	1.39	0.74
422	p1K14	0.87	0.94	1.02	1.06	0.96	0.95	1.39	1.74	1.21	0.96	1.02	1.08	1.14	1.56	1.10	1.30	1.52	0.63	0.53	0.42	1.80	1.69	1.89	0.74	1.04	0.66	0.65	0.58	0.64	1.02	0.93	1.04
424	p1K13	1.70	0.82	1.59	1.14	0.76	1.10	6.27	4.09	4.13	3.03	2.53	2.33	0.43	0.46	1.33	0.75	0.94	0.31	0.53	0.38	0.61	0.88	0.66	0.70	0.46	0.45	1.80	1.96	1.24	1.79	1.29	0.94
426	p1J20	1.07	0.84	1.33	0.87	1.05	0.29	1.50	1.73	1.35	1.30	1.44	0.95	1.27	1.84	0.87	0.58	0.58	1.10	0.48	0.15	0.75	0.87	1.10	1.47	0.97	0.73	1.16	0.80	0.63	1.40	1.13	0.79
428	p1J22	2.11	2.14	1.90	1.00	1.69	1.41	0.74	0.66	0.59	1.95	1.68	1.34	2.55	2.75	2.32	2.40	2.83	0.16	0.10	0.10	2.06	2.23	1.40	0.40	0.65	0.27	0.23	0.19	0.29	1.01	0.98	0.91
430	p1K1	1.09	1.27	0.65	1.74	0.79	0.67	1.18	1.46	1.08	0.47	0.46	0.56	1.06	1.93	1.11	0.70	1.03	1.03	4.33	3.72	0.40	0.66	0.58	1.23	1.33	1.38	1.30	0.74	1.01	0.88	0.96	0.65
432	p1K3	1.18	1.11	0.83	1.19	1.95	1.20	0.61	0.78	0.47	0.97	0.83	0.75	1.00	1.54	55.5	32.9	34.3	0.42	0.50	0.30	10.4	12.4	16.5	0.88	0.80	0.89	0.58	0.48	0.66	1.41	1.39	1.65
434	p1J19	1.62	1.24	0.85	1.63	2.03	2.49	0.50	0.73	0.54	0.76	0.58	0.68	0.99	1.77	7.90	4.54	4.19	0.76	0.95	0.62	21.6	18.8	17.4	0.69	0.73	0.62	0.41	0.38	0.60	2.16	2.58	1.92
434	p1K2	1.06	0.80	0.92	1.67	1.95	1.52	0.41	0.82	0.39	0.94	0.87	0.83	0.78	0.83	10.1	5.86	4.93	1.00	0.97	0.59	25.5	24.5	22.4	0.76	0.98	0.62	0.42	0.44	0.62	2.44	2.95	1.93
436	p1K5	0.93	0.43	0.81	0.93	0.54	0.54	1.58	1.27	1.28	1.24	0.89	1.19	0.23	0.31	0.17	0.18	0.17	1.11	1.34	0.98	0.19	0.27	0.38	2.68	2.64	2.54	1.49	1.76	1.46	1.27	1.08	1.18
438	p1J17	0.93	0.72	0.95	1.11	1.25	0.49	0.93	0.65	0.39	1.09	0.58	0.45	0.95	1.11	2.10	2.41	1.70	0.83	0.35	0.11	1.31	1.52	2.62	1.52	0.65	0.40	1.30	0.54	0.36	2.26	1.42	0.85
440	p1J18	0.31	0.18	0.36	1.10	0.53	0.26	1.04	0.44	0.24	0.94	0.43	0.44	1.35	1.30	2.83	4.02	2.20	1.47	0.49	0.16	1.57	2.24	3.13	2.02	0.64	0.41	1.76	0.77	0.52	3.09	2.01	0.82
442	p1J15	0.55	0.66	1.26	0.99	0.68	0.22	1.68	1.41	1.25	1.75	1.55	1.24	1.47	2.14	0.69	0.65	0.20	0.95	0.45	0.13	0.65	0.68	0.73	1.78	1.03	0.67	1.72	0.99	0.75	1.91	1.36	0.95
444	p1K4	0.41	0.21	0.23	1.61	1.07	1.06	1.38	1.29	0.96	0.90	0.65	0.68	0.67	1.07	0.66	0.35	0.30	2.98	2.04	0.80	0.10	0.11	0.18	1.14	0.83	1.71	1.43	1.81	2.39	1.96	3.28	3.23
446	p2A14	2.61	2.13	1.22	0.87	1.15	1.12	0.95	1.36	0.83	2.04	1.59	1.21	1.97	2.23	1.18	1.21	1.44	0.47	0.31	0.22	1.74	1.39	1.04	0.67	0.86	0.43	0.62	0.68	0.54	0.64	0.71	0.59

448	p1J23	0.96	0.57	0.50	1.05	1.18	1.38	0.73	1.01	0.70	0.54	0.42	0.54	0.94	1.33	10.5	5.19	2.21	21.1	25.3	28.6	9.69	14.3	8.43	0.79	0.84	0.66	0.40	0.55	0.42	1.10	1.12	0.93
450	p1J21	0.75	0.68	0.96	1.16	1.09	0.52	1.61	1.35	1.35	1.15	0.96	0.79	0.98	1.60	0.56	0.66	0.35	1.18	0.82	0.38	0.54	0.65	0.63	1.67	1.15	0.80	1.28	0.99	0.87	1.90	1.45	1.16
452	p1J24	1.62	0.56	0.84	1.98	1.51	0.96	1.56	2.04	1.37	1.05	0.81	0.52	0.56	0.62	0.30	0.27	0.28	1.05	0.90	0.60	0.21	0.24	0.25	2.13	1.32	0.57	1.95	1.15	0.76	2.32	2.16	1.11
454	p1J16	3.62	1.55	2.31	1.38	2.76	1.29	2.85	4.20	3.66	1.65	1.84	1.78	0.53	0.76	1.04	0.62	0.68	0.33	0.38	0.23	0.49	0.53	0.53	0.88	0.70	0.57	1.02	0.94	0.72	1.49	1.60	1.29
456	p1J2	0.77	0.76	0.37	1.57	0.75	0.51	2.52	3.12	2.29	0.38	0.28	0.45	0.72	1.14	0.63	0.30	0.43	1.28	1.69	0.86	0.57	0.52	0.42	2.02	2.14	1.12	1.16	2.18	1.75	1.62	1.56	0.75
458	p1J9	0.73	0.45	0.38	1.22	1.38	1.06	2.10	1.00	0.87	0.38	0.34	0.34	0.71	1.18	3.04	3.92	3.04	0.94	1.07	0.55	9.00	11.36	7.81	1.90	1.06	0.76	0.36	0.26	0.25	0.82	0.75	1.10
460	p1J10	0.86	0.48	0.72	1.18	1.10	0.85	1.70	2.20	1.82	1.20	0.92	1.08	0.26	0.32	0.62	0.48	0.39	0.78	0.97	0.64	1.48	1.10	0.84	1.34	0.95	0.78	1.34	1.56	1.29	1.40	1.82	1.25
462	p1J1	0.97	0.70	0.68	1.41	1.09	0.87	3.29	3.09	2.75	0.63	0.45	0.51	0.63	0.86	0.55	0.46	0.59	0.97	1.90	0.78	0.41	0.40	0.47	7.28	5.01	2.77	1.96	3.36	2.23	1.53	1.81	1.16
464	p1J5	2.26	0.94	0.74	56.2	22.5	16.6	1.12	2.88	2.58	1.22	0.79	0.83	1.06	1.15	3.02	1.07	1.06	0.53	0.98	0.47	0.61	0.97	0.55	0.50	0.63	0.48	1.09	0.78	1.24	1.19	0.95	1.04
466	p1J11	0.95	0.57	0.30	0.71	0.51	0.60	0.95	0.70	0.70	0.37	0.22	0.18	0.86	1.20	1.77	0.97	1.15	1.60	2.41	1.21	0.99	1.52	0.88	1.11	1.00	0.37	2.03	2.53	2.90	3.37	2.53	2.45
468	p1J8	1.77	0.67	1.20	1.73	0.98	0.86	1.16	1.88	1.28	1.70	1.59	1.57	0.28	0.34	0.11	0.08	0.06	0.55	0.87	1.22	0.08	0.14	0.14	3.15	2.84	1.11	2.76	2.77	2.38	0.89	0.97	0.64
470	p1J20	2.84	1.42	1.67	1.17	2.14	1.41	0.99	1.46	1.03	0.79	0.72	0.92	0.74	0.90	7.59	2.80	2.20	0.34	0.55	0.24	1.76	1.84	1.12	0.46	0.55	0.44	0.28	0.38	0.44	1.26	2.02	1.06
472	p1J3	0.70	0.39	0.64	1.75	1.28	1.04	1.24	1.61	1.81	0.94	0.69	0.65	0.33	0.52	0.53	0.51	0.65	0.99	0.72	2.08	0.28	0.37	0.41	2.76	3.52	1.80	1.06	1.71	1.56	2.20	2.39	1.64
474	p1J12	7.58	4.86	4.02	0.90	4.17	1.83	0.84	1.06	0.78	1.06	0.88	0.90	1.01	0.94	3.33	4.41	4.72	0.37	0.30	0.29	0.85	0.66	1.25	2.71	1.82	1.00	0.60	1.22	0.89	0.99	1.49	0.83
476	p1J23	1.35	0.87	0.71	1.36	2.00	1.43	0.71	1.03	0.70	0.80	1.12	1.23	0.41	0.84	2.67	1.69	3.03	0.46	0.77	0.41	1.20	15.2	26.0	0.53	0.59	0.38	0.51	0.49	0.49	1.33	1.69	1.01

478	p1j7	1.54	0.65	0.90	1.22	1.15	0.65	1.34	1.20	1.06	1.50	1.07	0.97	0.26	0.31	0.40	0.28	0.33	0.47	0.69	0.28	0.23	0.26	0.30	2.35	1.61	1.05	1.14	1.62	1.05	3.03	2.25	1.61
480	p1i21	1.76	0.96	0.81	31.8	11.1	9.16	11.33	2.83	2.80	1.14	0.85	0.92	1.16	1.74	1.19	0.42	0.57	0.43	0.63	0.30	0.53	0.65	0.44	0.63	0.78	0.60	1.19	1.06	1.14	1.24	1.12	1.06
482	p1i19	1.15	0.74	0.71	11.42	5.80	5.51	0.83	2.40	2.91	0.82	0.66	0.94	5.67	8.50	0.41	0.34	0.44	2.15	1.21	1.23	0.82	1.84	1.35	0.32	0.32	0.24	7.08	6.21	1.83	0.66	0.80	0.53
484	p1j4	1.92	0.80	0.63	2.15	0.85	0.63	3.22	7.24	4.65	2.55	1.77	1.05	0.91	1.08	0.13	0.07	0.08	0.55	0.55	0.31	0.13	0.18	0.18	5.44	2.85	3.01	5.06	4.87	1.88	1.09	1.73	0.77
486	p1i24	3.43	1.85	1.70	20.5	11.56	9.6	0.51	1.70	1.90	1.00	0.84	0.98	2.75	3.80	0.67	0.77	1.14	0.58	0.32	0.30	0.22	0.35	0.28	0.40	0.49	0.37	8.92	9.81	2.94	1.02	1.14	0.72
488	p1i18	1.27	0.84	0.91	6.72	5.15	4.95	1.01	1.60	1.28	0.81	0.81	0.67	0.80	1.49	0.85	0.58	0.79	0.65	0.75	0.81	1.44	2.23	1.77	0.99	1.12	0.98	0.69	1.24	1.55	2.56	2.75	1.99

TABLE 13 Response of Novel genes to Hypoxia

CLONE ID	GENE NAME	SEQ IDs	HIGHEST FOLD CHANGE IN HYPOXIA (hr hypoxia + cell type)
pIF6	Hypothetical protein hqp0376 protein	337/338	67.4 (18hr monocyte)
pIE7	Novel metallothionein	83/84	37.9 (18hr monocyte)
pID4	Hypothetical protein FLJ20500	25/26	23.8 (18hr monocyte)
pID1	Hypothetical protein FLJ10134	23/24	14.75 (18hr neuro)
pIH13	EST	193/194	12.5 DOWN (18hr mam epithelial)
pIF13	Hypothetical protein FLJ13356 fis, clone PLACE1000050	343/344	9.26 (18hr monocyte)
pIH6	EST	191/192	8.42 (6hr cardiomyocyte)
pIH17	EST	171/172	8.33 DOWN (18hr mam epithelial)
pIE14	unknown mRNA (schizophrenia-linked)	97/98	7.79 (6hr mam epithelial)
pIP14	Hypothetical protein KIAA1745	91/92	7.32 (18hr renal epithelial)
pIH19	EST	195/196	7.14 DOWN (18hr cardiomyocyte)
pID11	EST	135/136	6.90 (6hr mam epithelial)
pID17	Hypothetical protein KIAA1745	91/92	6.74 (6hr cardiomyocyte)
pIF9	Hypothetical protein KIAA0742	19/20	6.64 (18hr monocyte)
pID2	Hypothetical protein FLJ10134	23/24	6.62 (18 hr neuro)
pIH21	Hypothetical protein FLJ13511	163/164	6.61 (18hr monocyte)
pID16	cDNA FLJ20308 fis, clone HEP07264	33/34	6.29 (18hr neuroblastoma)
pID12	Hypothetical protein KIAA1376	29/30	5.98 (6hr cardiomyocyte)
pIH3	EST	215/216	5.88 DOWN (18hr mam epithelial)
pIH10	EST	189/190	4.98 (6hr cardiomyocyte)
pID18	cDNA FLJ13443 fis, clone PLACE1002853	127/128	4.84 (6hr cardiomyocyte)
pIH4	EST	213/214	4.76 DOWN (18hr mam epithelial)
pIH4	Hypothetical protein HSPC196	53/54	4.54 DOWN (18hr mam epithelial)
pIG20	cDNA YO23H03	203/204	4.17 DOWN (18hr cardiomyocyte)

p1115	Hypothetical protein CGI-117	47/48	4.17 DOWN (18hr adipocyte)
p1F8	Hypothetical protein KIAA0914	09-Oct	3.88 (6hr mam epithelial)
p1H20	EST	179/180	3.84 DOWN (18hr cardiomyocyte)
p1I22	Hypothetical protein KIAA1429	37/38	3.70 DOWN (18hr mam epithelial)
p1E16	cDNA DKFZp586E1624	65/66	3.56 (6hr endothelial)
p1H15	EST	177/178	3.45 DOWN (18hr mam epithelial)
p1F5	Hypothetical protein FLJ20281	11-Dec	3.36 (6hr mam epithelial)
p1E1	EST	123/124	3.24 (6hr cardiomyocyte)
p1H7	EST	175/176	3.13 DOWN (18hr mam epithelial)
p1D19	EST	143/144	2.98 (6hr cardiomyocyte)
p1F21	cDNA FLJ14342 fis, clone THYRO1000569	17/18	2.92 (18hr monocyte)
p1H12	EST	173/174	2.84 (6hr hepatocyte)
p1F2	Hypothetical protein FLJ20037	03-Apr	2.84 (18hr monocyte)
p1H23	cDNA FLJ21094 fis, clone CAS03807	187/188	2.78 DOWN (18hr neuroblastoma)
p1I13	Hypothetical protein FLJ11100	43/44	2.78 DOWN (18hr adipocyte)
p1D20	Hypothetical protein KIAA1125	139/140	2.73 (6hr renal epithelial)
p1E11	EST	109/110	2.73 (6hr hepatocyte)
p1D9	Hypothetical protein DKFZP564D116	27/28	2.71 (6hr adipocyte)
p1H5	Hypothetical protein FLJ22690	205/206	2.55 (6hr cardiomyocyte)
p1D24	EST	117/118	2.55 (18hr renal epithelial)
p1E12	Hypothetical protein DKFZP434E1723	69/70	2.49 (6hr mam epithelial)
p1F10	Hypothetical protein DKFZp434P0116	05-Jun	2.45 (6hr mam epithelial)
p1G22	EST	197/198	2.38 DOWN (18hr mam epithelial)
p1E4	EST	125/126	2.31 (18hr monocyte)
p1I3	Hypothetical protein FLJ11656	153/154	2.27 DOWN (18hr adipocyte)
p1F12	EST	01-Feb	2.27 (6hr hepatocyte)
p1G18	Mitochondrion sequence	211/212	2.17 (18hr neuroblastoma)
p1E23	cDNA FLJ14041 fis, clone HEMBA1005780	111/112	2.16 (18hr monocyte)
p1E9	novel PI-3-kinase adapter	79/80	2.15 (18hr monocyte)
p1G7	EST	281/282	2.14 (18hr monocyte)
p1J13	Hypothetical nuclear factor SBB122	35/36	2.13 DOWN (18hr adipocyte)
p1E19	EST	105/106	2.13 (6hr hepatocyte)

p1E15	cDNA Y127F12	107/108	2.12 (18hr monocyte)
p1F23	Hypothetical protein LOC51014	289/290	2.10 (6hr endothelial)
p2A14	EST	445/446	2.08 DOWN (18hr mam epithelial)
p1I17	Hypothetical protein FLJ20644	45/46	2.08 (6hr endothelial)
p1E8	cDNA: FLJ22249 fis, clone HRC02674	61/62	2.07 (6hr endothelial)
p1H1	Hypothetical protein FLJ10826	201/202	2.07 (18hr endothelial)
p1I10	EST	155/156	2.04 DOWN (18hr mam epithelial)
p1I5	Hypothetical protein FLJ10815	41/42	2.04 (6hr cardiomyocyte)
p1E20	Hypothetical protein FLJ20421	99/100	2.01 (6hr renal epithelial)
p1C23	cDNA FLJ12832 fis, clone NT2RP2003137	133/134	2.01 (6hr monocyte)
p1J16	cDNA: FLJ23019 fis, clone LNG00916	453/454	2.00 (6hr cardiomyocyte)
p1I12	Hypothetical protein MGC4549	151/152	1.96 DOWN (18hr mam epithelial)
p1F1	EST	81/82	1.96 (18hr monocyte)
p1E22	cDNA FLJ13618 fis, clone PLACE1010925	161/162	1.95 (6hr monocyte)
p1D21	Hypothetical protein FLJ22622	129/130	1.95 (6hr cardiomyocyte)
p1F19	Hypothetical protein KIAA0212	07-Aug	1.95 (18hr monocyte)
p1F18	Hypothetical protein KIAA0876	13/14	1.92 DOWN (6hr adipocyte)
p1H9	EST	185/186	1.91 (18hr monocyte)
p1F11	Hypothetical protein LOC51754	323/324	1.90 (6hr macrophage)
p1I2	cDNA FLJ11302 fis, clone PLACE1009971	149/150	1.88 (18hr monocyte)
p1G21	EST	199/200	1.85 DOWN (6hr macrophage)
p2A24	EST	101/102	1.85 DOWN (18hr adipocyte)
p1E13	Hypothetical protein PRO0823	21/22	1.85 DOWN (18hr adipocyte)
p1E10	cDNA FLJ11041 fis, clone PLACE1004405	71/72	1.84 (6hr endothelial)
p1I14	cDNA DKFZp564D016	147/148	1.80 (6hr monocyte)
p1J6	Hypothetical protein FLJ10206	39/40	1.78 (6hr mam epithelial)
p1F3	Hypothetical protein LOC94951	333/334	1.75 (6hr renal epithelial)
p1I16	Hypothetical protein KIAA1668	57/58	1.69 (6hr mam epithelial)
p1H16	EST	183/184	1.64 (18hr renal epithelial)
p1E17	cDNA FLJ31668 fis, clone NT2R12004916	103/104	1.63 (6hr cardiomyocyte)
p1I8	Hypothetical protein FLJ11296	55/56	1.61 (18hr macrophage)
p1H14	EST	167/168	1.45 (6hr renal epithelial)

TABLE 14 Response of Novel genes to Hypoxia

CLONE ID	GENE NAME	SEQ IDs	HIGHEST FOLD CHANGE IN HYPOXIA (hr hypoxia + cell type)
p1D6	ERO1 (S. cerevisiae)-like	67/68	11.30 (18hr fibroblast)
p1D10	Insulin induced protein 2	75/76	8.14 (18hr renal epithelial)
p1H2	Fatty acid binding protein 5	209/210	7.14 DOWN (18hr neuroblastoma)
p1H18	Ubiquitin specific protease 7	157/158	7.14 DOWN (18hr mam epithelial)
p1D22	MAX-interacting protein 1	119/120	6.68 (18hr renal epithelial)
p1C24	SLC25A19	93/94	6.13 (18hr macrophage)
p1E3	CYP1B1	137/138	5.88 DOWN (18hr renal epithelial)
p1G19	Mitochondrion sequence	207/208	5.88 DOWN (18hr mam epithelial)
p1D14	Clorf12	89/90	5.68 (6hr cardiomyocyte)
p1H8	ABL	181/182	4.76 DOWN (18hr cardiomyocyte)
p1E6	EGL nine (C.elegans) homolog 3	85/86	4.63 (18hr mam epithelial)
p1D13	A denylate kinase 3	77/78	4.58 (6hr cardiomyocyte)
p1H24	Nucleolar phosphoprotein Nopp34	159/160	4.40 (6hr cardiomyocyte)
p1D15	TRIP-Br2	31/32	4.09 (18hr renal epithelial)
p1F7	Spectrin, beta, non-erythrocytic 1	15/16	2.65 (6hr endothelial)
p1E5	Hepcidin antimicrobial peptide	141/142	2.59 (6hr macrophage)
p1C22	CD84-H1	131/132	2.58 (6hr cardiomyocyte)
p1E2	Mannosidase, alpha, class 1A, member 1	121/122	2.56 DOWN (18hr neuroblastoma)
p1C21	Tubulin, beta, 4	73/74	2.51 (6hr cardiomyocyte)
p1D3	Serine carboxypeptidase 1	95/96	2.49 (18hr fibroblast)
p1H11	Carboxypeptidase M	169/170	2.18 (18hr monocyte)
p1E18	Plexin C1	63/64	2.15 (6hr hepatocyte)
p2B1	PRAME	87/88	2.13 DOWN (18hr fibroblast)
p1E21	Glutamate-cysteine ligase, modifier subunit	113/114	2.04 (6hr hepatocyte)

p1111	SECIS binding protein 2	59/60	2.00 DOWN (18hr endothelial)
p111	Ribosomal RNA intergenic spacer	165/166	1.92 DOWN (18hr neuroblastoma)
p117	Uridine 5' monophosphate hydrolase 1	49/50	1.77 (18hr monocyte)
p1D23	PTEN	115/116	1.74 (6hr renal epithelial)
p1D5	ERO1 (S. cerevisiae)-like	67/68	1.72 (6hr mam epithelial)
p2A15	Sialyltransferase	145/146	1.61 DOWN (6hr monocyte)

TABLE 15 Genes with increased expression by macrophage activation

Clone	Seq ID	Gene Name	mRNA EXPRESSION (experimental condition)					
			#1	#2	#3	#4	#5	#6
p1K8	407/408	SCYA4	0.82	0.40	1.15	0.38	91.4	68.4
p1B16	251/252	Interleukin 8	0.75	1.13	0.47	0.41	42.8	28.1
p1B15	251/252	Interleukin 8	0.85	1.12	0.44	0.37	47.4	22.5
p1I21	479/480	SCYA8	0.54	0.18	1.15	0.32	19.6	12.2
p1I20	469/470	SCYA3L	0.92	0.41	1.00	0.30	29.4	22.8
p1N17	237/238	COX-2	0.90	1.00	0.84	0.84	18.9	20.3
p1J16	453/454	cDNA: FLJ23019 fis, clone LNG00916	0.92	0.66	0.91	1.15	14.4	14.9
p1I7	49/50	Uridine 5' monophosphate hydrolase 1	1.13	0.57	0.99	0.52	17.6	23.7
p1B14	251/252	Interleukin 8	0.71	1.20	0.51	0.47	10.1	21.4
p1E10	71/72	cDNA FLJ11041 fis, clone PLACE1004405	0.66	0.74	1.15	0.81	8.30	12.1
p2L23	397/398	endothelin 1	1.02	0.62	0.74	0.50	11.4	10.1
p1D19	143/144	EST	0.63	0.52	1.00	1.16	5.46	4.73
p1K3	431/432	Pleckstrin	1.14	0.70	0.73	0.54	6.49	2.34
p1C9	373/374	RAB-8b protein	0.95	0.81	0.77	0.94	5.11	4.53
p1I24	485/486	GRO1	0.90	0.72	0.78	1.04	4.69	2.96
p1G3	317/318	B-cell translocation gene 1	0.70	1.00	0.57	1.14	3.51	3.79
p1B1	243/244	Metallothionein 1G	0.51	1.00	0.66	1.85	2.50	3.83
p1J11	465/466	Fatty-acid-Coenzyme A ligase, long-chain 2	0.69	0.51	1.36	0.91	3.07	2.97
p1F17	329/330	P8 protein (candidate of metastasis 1)	0.26	1.78	0.16	0.88	1.16	2.59
p1F4	339/340	CYP1	0.60	1.04	0.77	1.15	2.52	4.22
p1D10	75/76	Insulin induced protein 2	0.49	1.00	0.48	1.23	1.63	4.96
p1E7	83/84	Novel metallothionein	0.49	1.26	0.70	1.11	1.32	2.89
p1D24	117/118	EST	0.58	0.71	1.00	1.32	1.56	2.59
p1I19	481/482	GRO2	0.99	1.00	0.69	0.55	2.65	2.29
p1E22	161/162	cDNA FLJ13618 fis, clone PLACE1010925	1.19	0.77	0.85	0.51	3.09	2.41
p1F6	337/338	Hypothetical protein hqp0376	0.44	1.08	0.47	1.06	1.11	2.47
p1J7	477/478	Sjogren syndrome antigen B	1.06	0.74	0.85	0.63	2.65	2.95
p1B19	235/236	plasminogen activator inhibitor, type 1	0.59	1.29	0.45	1.17	1.46	3.46
p1F24	297/298	Glia-derived nexin	0.65	0.68	0.99	1.10	1.62	1.80
p1P5	395/396	SCYA2	0.94	0.13	3.81	0.42	2.27	1.00
p1A22	263/264	Adenylate kinase 3	0.57	1.30	0.58	1.64	1.34	2.74
p1A23	265/266	Metallothionein 2A	0.55	0.89	0.95	1.01	1.23	4.38
p1A24	239/240	Metallothionein 1H	0.50	0.84	1.03	1.02	1.08	1.60

p1P3	221/222	PDGFB	0.47	2.06	0.31	1.99	1.00	1.49
p1A7	345/346	SLC31A2	1.10	0.92	0.84	0.93	2.30	2.32
p1P14	91/92/92a	Semaphorin 4b	0.47	2.52	0.95	4.04	0.96	3.88

Legend

mRNA expression values in the 6 experimental conditions (#1 no cytokines/ normoxia, #2 no cytokines/ hypoxia, #3 IL-10/ normoxia, #4 IL-10/ hypoxia, #5 LPS/IFN/ normoxia, #6 LPS/IFN/ hypoxia) are shown as values referenced to the median value of that gene throughout all 6 experimental conditions.

TABLE 16. Genes down-regulated by macrophage activation

Clone	Seq ID	Gene Name	mRNA EXPRESSION					
			(experimental condition)					
			#1	#2	#3	#4	#5	#6
p1H13	193/194	EST	1.35	1.41	1.00	0.91	0.44	0.68
p1E4	125/126	EST	1.22	1.13	1.09	0.96	0.40	0.40
p1G7	281/282	EST	1.30	1.44	1.01	1.51	0.54	0.82
p1E1	123/124	EST	1.21	1.64	0.94	1.35	0.51	0.63
p1D18	127/128	cDNA FLJ13443 fis, clone PLACE1002853	1.61	2.60	0.57	1.33	0.26	0.24
p1I2	149/150	cDNA FLJ11302 fis, clone PLACE1009971	2.39	1.23	1.07	0.54	0.45	0.43
p1G20	203/204	cDNA YO23H03	1.45	0.73	1.60	1.12	0.57	0.44
p1D21	129/130	Hypothetical protein FLJ22622	1.41	1.72	0.82	1.26	0.14	0.14
p1F8	9/10	Hypothetical protein KIAA0914	1.34	4.14	0.77	2.74	0.13	0.25
p1D16	33/34	Hypothetical protein FLJ20308	1.31	2.36	1.00	1.59	0.29	0.67
p1F3	333/334	Hypothetical protein XP_017131	1.63	2.21	0.92	1.00	0.42	0.42
p1D12	29/30	Hypothetical protein KIAA1376	0.89	2.62	0.79	2.07	0.28	2.61
p1I4	53/54	Hypothetical protein HSPC196	1.95	1.06	1.20	0.57	0.63	0.28
p1D9	27/28	Hypothetical protein DKFZP564D116	1.63	0.94	1.25	0.96	0.55	0.85
p1F9	19/20	Hypothetical protein KIAA0742	0.94	3.54	0.60	1.74	0.33	1.74
p1F11	323/324	Hypothetical protein LOC51754	1.67	1.91	1.00	0.86	0.60	0.59
p1I15	47/48	Hypothetical protein CGI-117	1.31	0.62	1.86	1.26	0.49	0.76
p1E13	21/22	Hypothetical protein PRO0823	1.15	0.93	1.15	1.08	0.44	0.24

p1F10	"5/6"	Hypothetical protein DKFZp434P0116	2.16	1.05	1.54	0.83	0.83	0.67
p1D1	23/24	Hypothetical protein FLJ10134	0.86	1.70	0.61	2.42	0.35	1.61
p1I5	41/42	Hypothetical protein FLJ10815	1.49	1.00	1.30	0.83	0.43	0.37
p1G13	293/294	ABCA1	1	1.05	1.22	1.07	0.43	0.46
p1B9	313/314	adipophilin	1.01	3.74	1.32	2.08	0.02	0.2
p1B7	313/314	adipophilin	1.57	3.44	0.77	1.17	0.21	0.45
p1B6	313/314	adipophilin	1.24	2.45	0.8	1.51	0.38	0.52
p1B8	313/314	adipophilin	1.11	1.87	1.01	1.04	0.55	0.59
p1K7	411/412	ATP-binding cassette E1	1.34	0.74	1.58	1.05	0.62	0.48
p1J23	447/448	Calgranulin A	1.21	0.94	3.35	2.8	0.58	0.86
p1K18	415/416	Colony-stimulating factor1	2.01	0.84	1.7	0.88	1	0.74
p1C2	331/332	CXCR4	1.01	3.76	0.46	1.47	0.08	1.13
p1C1	331/332	CXCR4	1.05	3.64	0.39	1.63	0.27	0.98
p1G12	321/322	Cyclin G2	0.85	2.17	0.6	1.29	0.28	1.33
p1F16	325/326	CYP1B1	1.37	0.96	1.67	0.66	0.64	0.48
p1C7	369/370	D123	1.69	1.33	1.1	0.7	0.65	0.83
p1G17	315/316	Early development regulator 2	0.97	2.47	1.12	2.24	0.29	0.85
p1I23	475/476	Ecotropic viral integration site 2A	1.39	1.25	1.11	1.72	0.18	0.22
p1A14	257/258	Enolase 1	0.99	3.22	1.19	2.46	0.13	0.37
p1A10	273/274	Enolase 2	1.17	5.28	0.59	3.77	0.49	1.08
p1D6	67/68	ERO1 (S. cerevisiae)-like	0.84	3.02	0.97	2.87	0.32	1.58
p1A11	253/254	GAPDH	1.21	2.41	0.93	1.31	0.32	0.81
p1A12	253/254	GAPDH	1.09	1.97	1	1.49	0.41	0.96
p1K22	419/420	GPR44	1.24	1.03	1.42	0.93	0.56	0.48
p1C18	269/270	Granulin	1.28	1.59	0.96	1.03	0.56	0.72
p1C17	269/270	Granulin	1.58	1.6	0.62	0.41	0.76	0.94
p1A15	249/250	Hexokinase-2	0.89	3.88	0.68	3.11	0.38	2.02
p1C13	381/382	Jk-recombination signal binding protein	1.11	1.18	1.43	1.98	0.32	0.73
p1A8	223/224	Lactate dehydrogenase A	0.7	2.25	1.4	1.44	0.26	1.32
p1A9	223/224	Lactate dehydrogenase A	0.77	1.85	1.15	1.68	0.32	1.19
p1G5	279/280	MAX-interacting protein 1	1.24	5.5	0.9	4.48	0.34	0.97
p1D22	119/120	MAX-interacting protein 1	1.2	3.86	0.52	3.44	0.37	0.91
p1G18	211/212	Mitochondrion sequence	1.27	1.12	1	1.31	0.57	0.77
p1K23	403/404	MYC	1.37	0.77	2.39	1.09	0.54	0.35
p1E20	99/100	Myo-inositol monophosphatase A3	1.12	1.28	1.02	0.99	0.48	0.61
p1B20	267/268	Osteopontin	1.13	1.58	0.99	1.52	0.1	0.4
p1F13	343/344	Papillomavirus regulatory factor PRF-1	0.98	5.02	0.44	6.79	0.09	2.43
p1A13	255/256	Phosphoglycerate kinase 1	1.04	2.45	1.23	1.83	0.2	0.9
p1G9	305/306	PI-3-kinase, catalytic, beta polypeptide	1.46	1.88	0.75	1.17	0.44	0.47

p1E18	63/64	Plexin C1	1.72	1.79	1	0.85	0.69	0.35
p1C11	377/378	polyubiquitin	1.13	1.79	0.79	1.14	0.5	0.84
p1B3	231/232	Proline 4-hydroxylase, alpha polypeptide I	0.94	1.38	1.03	1.58	0.43	0.89
p1B4	349/350	Proline 4-hydroxylase, alpha polypeptide II	0.9	1.46	1.05	1.41	0.44	1
p1B22	355/356	Protease, serine, 11	1.3	1.1	1.26	0.92	0.64	0.7
p1C10	375/376	Regulator of G-protein signalling 1	1.42	1.68	0.94	1.55	0.47	0.95
p1D3	95/96	Serine carboxypeptidase 1	1.22	1.07	1.07	1.09	0.33	0.88
p1F15	341/342	SHB adaptor protein	1.04	1.61	0.94	1.72	0.43	0.54
p1A5	311/312	SLC2A5	0.71	2.6	1.06	2.09	0.34	1.09
p1G4	307/308	SLC5A3	1.12	1.44	0.93	1.31	0.33	0.62
p1A20	261/262	Triosephosphate isomerase 1	0.97	2.06	1.09	2.24	0.17	0.66
p1D15	31/32	TRIP-Br2	1.16	1.4	1.1	1.25	0.47	0.46
p1K4	443/444	TSC-22	1.44	1	1.55	0.7	0.6	0.57

TABLE 17: Genes responsive to IL-10 (increased or decreased) but not affected significantly by LPS+IFN

Clone	Seq ID	Gene Name	mRNA EXPRESSION					
			(experimental condition)					
			#1	#2	#3	#4	#5	#6
p1H8	181/182	ABL	1.02	0.96	6.65	5.25	0.86	0.73
p1E15	107/108	cDNA Y127F12	0.48	0.77	1.69	2.45	0.78	1.44
p2A14	445/446	EST	1.06	0.74	2.78	3.09	1.06	0.82
p1H6	191/192	EST	1.01	0.84	2.47	2.05	0.93	0.83
p1E5	141/142	Hepcidin antimicrobial peptide	0.84	0.73	1.91	1.68	0.58	2.16
p1H12	151/152	Hypothetical protein MGC4549	1.07	0.67	2.34	2.53	1.11	0.74
p1D8	271/272	Hypoxia-inducible protein 2	0.65	1.00	1.51	1.89	0.71	2.05
p1K14	421/422	Keratin 6B	1.03	0.68	3.80	3.28	0.97	0.76
p1J22	427/428	Neutral sphingomyelinase (N-SMase) activation associated factor	0.94	0.79	5.59	3.52	0.91	1.29
p1G15	287/288	Phosphoglucomutase 1	0.82	1.20	1.83	1.90	0.61	1.05
p1A2	247/248	SLC2A3	0.37	3.31	1.00	3.32	0.49	2.63
p1A3	247/248	SLC2A3	0.39	2.45	1.00	2.65	0.20	1.50
p1K2	433/434	CFFM4	1.30	0.98	0.51	0.59	1.11	0.91
p1C4	363/364	FGF receptor activating protein 1	1.02	0.96	0.50	0.63	1.16	1.31

TABLE 18. Genes up-regulated in human tumors. Individual patients are denoted by the letters E,F,G,H and K.

Clone	Gene Name	SeqID	Ovary nor	Ovary tum	Ovary nor	Ovary tum	Ovary nor	Ovary tum	Ovary nor	Ovary tum	Breast nor	Breast tum	Breast nor	Breast tum
p1H8	ABL	182	0.73	2.21	0.72	1.48	1.15	3.15	2.41	2.01	2.61	1.00		
p1B6	adipophilin	314	0.44	1.57	0.51	0.37	1.30	0.99	0.58	0.78	0.82	1.06		
p1A19	Aldolase C	260	0.27	1.00	0.74	1.06	0.40	1.49	0.62	0.47	0.49	2.18		
p1C2	CXCR4	332	0.29	0.91	1.03	1.41	2.43	2.80	2.71	0.95	1.81	0.59		
p1K1	Cyclophilin F	430	0.60	0.71	1.11	0.80	0.61	1.85	0.76	0.76	0.95	1.66		
p1E3	CYP1B1	138	0.32	0.06	0.45	1.47	1.05	0.16	1.00	0.38	1.20	0.16		
p1F16	CYP1B1	326	0.60	0.17	0.67	2.30	1.65	0.24	1.00	0.55	1.82	0.24		
p1C8	Decl	372	0.93	0.66	1.85	1.10	1.37	0.56	0.94	0.53	0.87	5.55		
p1A14	Enolase 1	258	0.10	0.46	1.00	1.26	0.41	1.47	0.36	0.81	0.53	0.61		
p1A10	Enolase 2	274	0.63	0.64	1.48	3.23	0.80	2.67	0.96	0.92	0.92	0.52		
p1D6	ERO1 (S. cerevisiae)-like	68	0.30	1.61	0.73	1.32	0.77	0.31	0.66	0.42	1.10	1.02		
p1E19	EST	106	1.00	2.04	1.38	1.30	0.20	1.19	2.38	2.49	1.19	1.78		
p1H15	EST	178	0.22	0.46	0.71	1.07	0.82	2.57	2.62	1.22	3.66	1.07		
p1H16	EST	184	1.57	2.66	1.44	1.42	0.74	2.65	2.32	1.91	1.70	1.95		
p1H17	EST	172	0.65	2.20	1.00	1.62	0.95	2.64	2.95	2.32	2.98	0.93		
p1H20	EST	180	0.81	2.39	0.98	1.61	0.78	2.91	2.83	1.62	2.64	1.17		
p1H3	EST	216	0.10	0.38	0.78	1.14	0.88	2.49	2.26	1.18	3.08	0.89		
p1C4	FGF receptor activating protein 1	364	0.53	0.77	0.91	1.00	1.04	0.98	0.81	1.27	0.72	2.43		
p1A11	GAPDH	254	0.04	0.74	1.50	2.84	0.86	1.57	0.27	1.12	0.51	0.69		
p1C6	Glucose phosphate isomerase	368	0.21	0.86	1.59	2.26	0.65	1.94	0.47	0.79	0.61	1.00		
p1E21	Glutamate-cysteine ligase, modifier subunit	114	1.26	2.38	1.35	1.23	0.27	1.51	2.05	4.12	1.00	1.78		
p1C18	Granulin	270	0.44	0.91	1.00	0.83	0.71	0.88	0.60	0.61	0.73	2.72		

p1D21	Hypothetical protein FLJ22622	130	0.39	8.14	0.53	0.59	0.92	0.54	0.87	1.71	0.99	1.00
p1F11	Hypothetical protein LOC51754	324	0.53	0.77	1.01	0.97	0.61	1.88	1.54	1.10	0.85	0.70
p1A8	Lactate dehydrogenase A	224	0.24	1.07	1.09	1.00	0.70	0.40	0.52	0.50	0.66	0.92
p1K9	Lipocortin I	402	0.46	4.09	1.33	1.17	1.07	0.31	1.20	0.54	0.23	0.12
p1B1	Metallothionein IG	244	1.21	0.60	3.14	1.49	1.00	2.04	0.46	1.54	0.27	0.51
p1K23	MYC	404	5.98	1.00	2.50	2.41	4.11	1.91	0.64	2.17	0.32	0.44
p1B20	Osteopontin	268	0.05	1.01	2.74	1.69	0.22	0.36	0.24	0.52	0.35	0.92
p1B21	Osteopontin	268	0.22	1.73	2.09	1.87	0.58	0.76	0.52	0.65	0.44	1.00
p1F17	P8 protein (candidate of metastasis I)	330	1.18	0.17	1.27	1.47	0.55	1.79	1.00	0.97	0.92	1.33
p1C11	polyubiquitin	378	1.11	0.91	1.59	1.53	1.00	1.34	0.63	0.65	0.51	2.52
p2B1	PRAME	88	1.13	8.86	5.08	9.57	1.15	18.20	1.63	1.64	0.96	0.96
p1B5	Proline 4-hydroxylase, alpha polypeptide II	350	0.50	0.52	1.02	3.18	1.94	1.31	0.79	0.89	0.72	1.62
p1P14	Semaphorin 4b	92	0.80	4.26	0.86	1.01	1.22	1.88	1.33	1.73	1.48	2.26
p1A6	SLC2A5	312	0.52	5.07	0.89	2.12	1.48	0.68	1.21	0.55	1.37	0.87
p1J17	SLC6A1	438	0.11	0.39	0.85	1.42	0.80	3.56	3.08	1.17	4.08	1.00
p1J18	Synaptopodin	440	0.07	0.31	0.99	1.44	0.68	3.03	3.51	1.40	4.54	1.29
p1J15	TERA protein	442	0.68	2.14	0.80	1.47	1.00	2.36	2.97	2.74	3.14	0.97
p1G11	Tumor protein D52	302	0.20	1.49	0.51	1.07	0.83	1.00	1.66	1.82	1.90	2.62
p1H18	Ubiquitin specific protease 7	158	0.73	2.13	0.91	1.76	0.82	2.47	2.56	2.04	2.53	1.00
p1O20	VEGF	228	0.84	4.19	0.85	1.55	2.31	12.69	0.66	1.42	1.35	6.44

TABLE 19. Genes down-regulated in human tumors. Individual patients are denoted by the letters E,F,G,H and K.

Clone	Gene Name	Seq1	Ovary nor	Ovary tum	Ovary nor	Ovary tum	Ovary nor	Ovary tum	Breast nor	Breast tum	Breast nor	Breast tum
pIC3	Activin A receptor, type I	362	1.51	0.32	1.28	1.76	1.06	1.21	0.99	0.87	0.76	1.17
pIB9	adipophilin	314	0.67	0.20	0.78	0.50	1.53	2.45	0.61	0.83	0.68	0.85
pIK15	Alpha-2-macroglobulin	406	0.39	0.12	0.79	0.54	1.08	0.38	0.71	1.16	0.53	0.71
pIG3	B-cell translocation gene 1	318	2.01	0.64	1.15	1.13	1.69	0.58	1.39	1.13	1.66	0.78
pIF14	Butyrate response factor 1	328	2.85	0.94	1.86	2.31	1.46	0.79	1.36	1.00	0.86	1.09
pIJ23	Calgranulin A	448	0.43	0.90	1.00	0.91	1.60	0.59	12.02	1.60	23.70	0.80
pIK2	CFFM4	434	0.47	0.29	1.30	1.06	2.32	0.31	1.00	0.49	0.79	1.52
pIJ19	CFFM4	434	0.45	0.29	1.24	1.00	2.06	0.35	1.42	0.56	0.76	1.44
p2A23	Chitinase 3-like 2	284	0.66	0.78	0.48	0.61	4.18	0.74	1.36	2.22	0.91	2.01
pIN17	COX-2	238	0.73	1.21	0.51	0.72	2.31	0.57	0.80	0.61	0.56	0.54
pIC2	CXCR4	332	0.29	0.91	1.03	1.41	2.43	2.80	2.71	0.95	1.81	0.59
pIE3	CYP1B1	138	0.32	0.06	0.45	1.47	1.05	0.16	1.00	0.38	1.20	0.16
pIF16	CYP1B1	326	0.60	0.17	0.67	2.30	1.65	0.24	1.00	0.55	1.82	0.24
pIC8	Decl	372	0.93	0.66	1.85	1.10	1.37	0.56	0.94	0.53	0.87	5.55
pIJ10	DNCL12	460	1.00	0.98	1.28	1.17	1.66	0.50	2.02	1.51	0.98	0.90
Ecotropic viral integration site												
pIJ23	2A	476	0.64	0.67	0.88	1.08	1.28	0.39	0.79	0.71	1.06	0.78
pIE4	EST	126	1.30	1.67	0.88	0.65	0.70	0.17	1.00	0.69	1.06	0.46
pIH19	EST	196	0.71	1.52	0.80	0.99	1.08	1.17	1.99	1.70	3.34	0.83
pIH4	EST	214	0.32	0.76	0.70	1.02	1.14	2.36	2.34	1.65	2.90	0.75
pIH3	EST	216	0.10	0.38	0.78	1.14	0.88	2.49	2.26	1.18	3.08	0.89
pIH15	EST	178	0.22	0.46	0.71	1.07	0.82	2.57	2.62	1.22	3.66	1.07
pIH17	EST	172	0.65	2.20	1.00	1.62	0.95	2.64	2.95	2.32	2.98	0.93
Fatty-acid-Coenzyme A												
pIJ11	ligase, long-chain 2	466	0.83	0.23	0.68	1.38	1.05	0.44	1.42	0.85	1.00	1.00

p1F24	Glia-derived nexin	298	3.17	0.93	3.24	0.85	1.26	0.80	1.64	1.89	1.25	0.96
p1I24	GRO1	486	0.82	2.23	1.55	3.98	3.19	0.52	0.50	0.55	0.31	0.52
p1I19	GRO2	482	1.97	2.36	3.67	4.14	11.31	1.00	1.29	1.48	0.64	0.70
p1D1	Hypothetical protein	24	1.20	0.40	1.82	1.82	0.73	0.68	0.81	1.03	1.28	0.96
p1H21	Hypothetical protein	164	1.00	0.22	0.97	0.49	1.73	0.72	1.35	0.75	2.85	1.76
p1F2	Hypothetical protein	4	2.41	0.61	1.20	0.63	1.40	0.39	0.66	0.73	0.53	0.63
p1F23	Hypothetical protein	290	1.27	0.85	0.90	1.03	1.82	0.55	0.90	0.92	0.84	0.94
p1E13	Hypothetical protein	22	0.51	0.60	1.00	0.98	1.74	0.52	2.15	0.98	2.26	0.85
p1F3	Hypothetical protein	334	2.75	0.91	1.26	1.39	1.42	0.61	1.67	1.96	1.13	1.02
p1B14	Interleukin 8	252	2.96	0.36	4.71	1.04	8.45	1.00	0.36	0.37	0.54	0.37
p1B16	Interleukin 8	252	3.16	0.50	7.46	1.82	16.49	3.29	1.00	0.42	0.66	0.36
p1B15	Interleukin 8	252	3.43	0.54	7.25	1.69	11.81	3.46	0.94	0.75	0.67	0.88
p1C13	Jk-recombination binding protein	382	1.94	0.55	1.48	1.17	0.54	0.50	0.54	0.80	0.51	0.57
p1K9	Lipocortin I	402	0.46	4.09	1.33	1.17	1.07	0.31	1.20	0.54	0.23	0.12
p1E2	Mannosidase, alpha, class 1A, member 1	122	0.87	0.40	1.00	0.88	2.02	0.50	1.51	0.67	1.54	1.63
p1A23	Metallothionein 2A	266	0.78	0.22	2.96	1.22	2.34	1.05	0.58	1.50	0.40	0.48
p1G19	Mitochondrion sequence	208	1.00	1.74	1.17	1.08	0.99	1.69	2.16	1.67	2.10	0.58
p1G18	Mitochondrion sequence	212	0.77	1.43	0.95	1.02	1.00	0.99	1.71	1.62	2.61	0.86
p1K23	MYC	404	5.98	1.00	2.50	2.41	4.11	1.91	0.64	2.17	0.32	0.44
p1J20	Neuro-oncological antigen 1	426	0.54	1.61	1.00	1.50	1.24	2.65	3.08	2.55	2.77	0.83
p1F17	P8 protein (candidate of metastasis 1)	330	1.18	0.17	1.27	1.47	0.55	1.79	1.00	0.97	0.92	1.33
p1B19	plasminogen activator	236	3.22	0.44	2.15	1.84	7.67	2.48	0.97	0.56	0.88	0.86

Inhibitor, type I	plasminogen activator	236	1.53	0.34	1.27	1.28	2.90	1.44	0.52	0.84	0.79	0.75
p1B18	Inhibitor, type I	432	0.39	0.45	0.99	0.92	1.78	0.43	1.39	0.41	1.66	1.00
p1K3	Pleckstrin											
	Proline 4-hydroxylase, alpha	232	1.12	1.00	1.44	1.83	0.71	0.74	0.78	0.25	1.22	0.50
p1B3	polypeptide I											
	Proline-rich protein with											
p1F20	nuclear targeting signal	336	4.29	0.54	2.08	1.23	2.34	1.08	0.89	1.00	1.13	0.61
p1B22	Protease, serine, II	356	3.86	0.86	9.76	2.40	1.17	0.95	1.18	0.88	0.79	1.00
	Regulator of G-protein											
p1C10	signalling I	376	0.26	0.12	1.09	1.10	1.93	0.14	0.42	0.12	0.44	1.00
p1P5	SCYA2	396	2.43	0.63	3.14	1.39	2.20	0.61	1.14	0.98	0.82	0.89
p1K8	SCYA4	408	1.00	0.56	2.48	2.29	1.94	0.54	1.94	0.68	1.75	1.62
p1I11	SECIS binding protein 2	60	1.49	0.90	1.00	0.98	1.09	0.36	2.31	2.24	2.10	1.49
p1I18	Selectin L	488	0.60	0.78	1.00	1.46	4.25	1.46	5.73	1.09	4.30	2.26
p1D3	Serine carboxypeptidase I	96	0.76	0.09	0.91	0.98	0.76	0.50	0.98	1.00	1.13	0.95
p1I17	SLC6A1	438	0.11	0.39	0.85	1.42	0.80	3.56	3.08	1.17	4.08	1.00
	Spectrin, beta, non-											
p1F7	erythrocytic I	16	2.52	0.48	2.04	1.14	1.65	0.95	2.30	1.90	0.91	0.91
p1J18	Synaptopodin	440	0.07	0.31	0.99	1.44	0.68	3.03	3.51	1.40	4.54	1.29
p1J15	TERA protein	442	0.68	2.14	0.80	1.47	1.00	2.36	2.97	2.74	3.14	0.97
p1K4	TSC-22	444	2.92	0.46	1.40	2.19	0.85	1.05	0.56	1.12	0.35	0.88

TABLE 20: Genes up-regulated in response to TNF α

Clone	Gene Name	SeqID	Cytokine / % Oxygen			
			none 20	none 0.1	TNF 20	TNF 0.1
p1C14	Abstrakt	384	1	7.89	4.63	7.87
p1D13	Adenylate kinase 3	78	1	0.85	2.36	2.47
p1A22	Adenylate kinase 3	264	1	1.53	2.69	3.72
p1B8	adipophilin	314	1	17.1	2.27	8.81
p1B7	adipophilin	314	1	13.4	2.17	5.86
p1A19	Aldolase C	260	1	6.61	2.57	6.31
p1N17	COX-2	238	1	1.04	0.91	2.24
p1C1	CXCR4	332	1	5.42	2.21	5.22
p1F4	CYP1	340	1	2.86	3.43	6.26
p1E3	CYP1B1	138	1	0.33	2.12	1.14
p1F16	CYP1B1	326	1	0.45	1.93	1.19
p2L23	endothelin 1	398	1	0.95	2.74	2.41
p1A14	Enolase 1	258	1	9.98	7.22	11.78
p1A11	GAPDH	254	1	7.87	3.60	5.90
p1C6	Glucose phosphate isomerase	368	1	5.18	2.58	3.61
p1D9	Hypothetical protein DKFZP564D116	28	1	2.37	2.48	3.16
p1F5	Hypothetical protein FLJ20281	12	1	3.84	2.05	3.78
p1B23	interleukin 1 receptor antagonist	358	1	3.46	3.43	5.35
p1B14	Interleukin 8	252	1	5.52	16.8	56.8
p1B16	Interleukin 8	252	1	2.55	9.64	23.3
p1B15	Interleukin 8	252	1	3.37	10.4	28.1
p1C13	Jk-recombination signal binding protein	382	1	5.82	4.77	8.75
p1A8	Lactate dehydrogenase A	224	1	24.8	4.08	15.1
p1A13	Phosphoglycerate kinase 1	256	1	7.29	2.73	4.65
p1B19	Plasminogen activator inhibitor, type 1	236	1	3.78	2.63	9.41
p1B18	Plasminogen activator inhibitor, type 1	236	1	4.92	2.23	6.55
p1C11	Polyubiquitin	378	1	2.80	2.06	3.03
p1B4	Proline 4-hydroxylase, alpha polypeptide II	350	1	6.15	3.09	5.80
p1F20	Proline-rich protein with nuclear targeting signal	336	1	4.69	2.18	6.45
p1I20	SCYA3L	470	1	0.77	3.97	3.61
p1K8	SCYA4	408	1	0.81	9.65	9.63
p1D3	Serine carboxypeptidase 1	96	1	3.74	2.37	3.55
p1A2	SLC2A3	248	1	16.0	2.68	15.5
p1F22	Sorting nexin 9	320	1	0.66	1.26	1.63
p1B10	Stearoyl-CoA desaturase	352	1	5.04	3.05	6.95

241

p1B17	Tissue factor	226	1	3.69	2.74	6.03
p1A20	Triosephosphate isomerase I	262	1	16.1	8.30	16.2
p1E14	unknown mRNA (schizophrenia-linked)	98	1	3.30	3.03	3.26

TABLE 21: Genes down-regulated in response to TNF α

Clone	Gene Name	SeqID	Cytokine / % Oxygen			
			none 20	none 0.1	TNF 20	TNF 0.1
p1E5	Hepcidin antimicrobial peptide	142	1	1.50	0.19	0.70
p1H2	Fatty acid binding protein 5	210	1	0.72	0.38	0.46
p1P5	SCYA2	396	1	0.29	0.45	0.26
p1J5	SCYA7	464	1	0.89	0.46	0.49

5

TABLE 22: Genes up-regulated in response to IL-17

Clone	Gene Name	SeqID	Cytokine / % Oxygen			
			none 20	none 0.1	IL-17 20	IL-17 0.1
p1J11	Fatty-acid-Coenzyme A ligase, long-chain 2	466	1	0.55	1.63	1.28
p1A11	GAPDH	254	1	0.78	2.60	1.84
p1C6	Glucose phosphate isomerase	368	1	0.84	2.14	1.43
p1I24	GRO1	486	1	1.02	2.29	1.28
p1I19	GRO2	482	1	1.02	2.26	1.43
p1B16	Interleukin 8	252	1	1.77	9.52	12.2
p1B15	Interleukin 8	252	1	1.54	7.36	9.71
p1B14	Interleukin 8	252	1	1.50	9.34	7.13
p1P5	SCYA2	396	1	0.24	2.12	0.58
p1K8	SCYA4	408	1	0.44	2.48	0.83

TABLE 23: Genes down-regulated in response to IL-17

Clone	Gene Name	SeqID	Cytokine / % Oxygen			
			none	none	IL-17	IL-17
			20	0.1	20	0.1
p1H8	ABL	182	1	1.08	0.08	0.09
p1J22	Neutral sphingomyelinase (N-SMase) activation associated factor	428	1	1.21	0.13	0.11
p1K14	Keratin 6B	422	1	1.27	0.15	0.15
p1J6	Hypothetical protein FLJ10206	40	1	1.28	0.22	0.23
p1I12	Hypothetical protein MGC4549	152	1	1.20	0.32	0.30
p1E15	cDNA Y127F12	108	1	1.58	0.21	0.56
p2A14	EST	446	1	1.10	0.56	0.40
p1G24	Glycogen synthase 1	276	1	1.45	0.34	0.65
p1C16	Decidual protein induced by progesterone	388	1	1.09	0.73	0.51
p1D8	Hypoxia-inducible protein 2	272	1	1.30	0.44	0.65
p1B18	Plasminogen activator inhibitor, type 1	236	1	1.10	0.49	0.78
p1H4	EST	214	1	1.13	0.49	0.58

TABLE 24: Genes up-regulated in response to IL-15

Clone	Gene Name	SeqID	Cytokine / % Oxygen			
			none	none	IL-15	IL-15
			20	0.1	20	0.1
p1A19	Aldolase C	260	1	0.50	0.35	1.30
p1J16	cDNA: FLJ23019 fis, clone LNG00916	454	1	0.80	5.76	7.27
p1D19	EST	144	1	1.00	2.27	1.39
p1J11	Fatty-acid-Coenzyme A ligase, long-chain 2	466	1	0.55	1.64	1.29
p1A11	GAPDH	254	1	0.78	0.52	4.32
p1C6	Glucose phosphate isomerase	368	1	0.84	0.57	3.13
p1H5	Hypothetical protein FLJ22690	206	1	0.94	2.37	1.59
p1A23	Metallothionein 2A	266	1	1.41	1.26	3.08
p1P5	SCYA2	396	1	0.24	4.51	1.37
p1J5	SCYA7	464	1	0.66	3.27	1.61
p1I21	SCYA8	480	1	0.37	3.77	1.55
p1I7	Uridine 5' monophosphate hydrolase 1	50	1	0.84	4.98	3.61

TABLE 25: Genes down-regulated in response to IL-15

Clone	Gene Name	SeqID	Cytokine / % Oxygen			
			none 20	none 0.1	IL-15 20	IL-15 0.1
p1H8	ABL	182	1	1.08	1.22	0.09
p1C14	Abstrakt	384	1	0.69	0.40	1.24
p1B8	Adipophilin	314	1	1.08	0.23	1.42
p1B7	Adipophilin	314	1	1.10	0.30	2.02
p1B6	Adipophilin	314	1	0.98	0.37	2.39
p1B9	Adipophilin	314	1	1.46	0.41	1.66
p1A19	Aldolase C	260	1	0.50	0.35	1.30
p1C7	D123	370	1	0.53	0.47	0.90
p1H6	EST	192	1	1.46	1.95	0.67
p2A14	EST	446	1	1.10	1.08	0.51
p1G24	Glycogen synthase 1	276	1	1.45	0.88	0.55
p1J6	Hypothetical protein FLJ10206	40	1	1.28	1.28	0.18
p1I12	Hypothetical protein MGC4549	152	1	1.20	1.64	0.28
p1D8	Hypoxia-inducible protein 2	272	1	1.30	1.25	0.62
p1K14	Keratin 6B	422	1	1.27	1.48	0.11
p1A8	Lactate dehydrogenase A	224	1	1.31	0.48	2.29
p1A9	Lactate dehydrogenase A	224	1	1.95	0.49	1.83
p1J22	Neutral sphingomyelinase (N- SMase) activation associated factor	428	1	1.21	1.51	0.11
p1B4	Proline 4-hydroxylase, alpha polypeptide II	350	1	0.81	0.50	1.34
p1A20	Triosephosphate isomerase 1	262	1	0.81	0.36	1.22

TABLE 26 cross-references all protein and nucleotide sequences (SEQ ID Nos.) that are referenced herein to accession numbers in public databases available as of 8.12.00.

TITLE	Hypoxia response	PROTEIN		NUCLEOTIDE	
		SEQ ID	ACCESSION	SEQ ID	ACCESSION
cDNA FLJ13611 fis, clone PLACE1010802	Increase	1	BAB14633	2	AK023673
Hypothetical protein FLJ20037	Increase	3	BAA90903	4	AK000044
hypothetical protein DKFZp434P0116	Increase	5	CAB70863	6	AL137661
KIAA0212	Increase	7	BAA13203	8	D86967
KIAA0914	Increase	9	BAA74937	10	AB020721
Hypothetical protein FLJ20281	Increase	11	NP_060212	12	NM_017742
KIAA0876	Increase	13	BAA74899	14	AB020683
cDNA FLJ13700 fis, clone PLACE2000216	Increase	15	(nearest=Q01082)	16	AK023762
DKFZP586G1122 protein	Increase	17	CAB55938	18	AL117462
Putative zinc finger protein LOC55818	Increase	19	AAF67005	20	AF153648
hypothetical protein PRO0823	Increase	21	AAF71073	22	AF116653
Hypothetical protein FLJ10134	Increase	23	BAA91458	24	AK000996
Hypothetical protein FLJ20500	Increase	25	BAA91214	26	AK000507
DKFZP564D116 protein	Increase	27	CAB43242	28	AL050022
KIAA1376 protein	Increase	29	BAA92614	30	AB037797
Hypothetical protein KIAA0127	Increase	31	BAA09476	32	D50917
Hypothetical protein FLJ20308	Increase	33	BAA91078	34	AL137263
Hypothetical nuclear factor SBB122	Repression	35	NP_065128	36	NM_020395
DKFZP434I116 protein	Repression	37	CAB55922	38	AL117434
Hypothetical prot. FLJ10206	Repression	39	NP_060495	40	NM_018025
hypothetical protein FLJ10815	Repression	41	BAA91830	42	AK001677
Hypothetical protein FLJ11100	Repression	43	BAA92003	44	AK001962

hypothetical protein FLJ2064	Repression	45	NP_060387	46	NM_017917
Hypothetical protein HSPC111	Repression	47	NP_057475	48	NM_016391
hypothetical protein LOC51251	Repression	49	NP_057573	50	NM_016489
KIAA0014	Repression	51	BAA04946	52	D25216
Hypothetical protein HSPC196	Repression	53	NP_057548	54	NM_016464
Hypothetical protein FLJ11296	Repression	55	BAA92115	56	AK002158
Hypothetical protein bA395L14	Repression	57	CAB62980	58	AL022311
cDNA FLJ13016 fis, clone NT2RP3000624	Repression	59	BAB14393	60	AK023078
cDNA DKFZp586H0324 clone DKFZp586H0324	Increase	61	none	62	AL110163
Clone 23785	Increase	63	none	64	AF035307
cDNA DKFZp586E1624	Increase	65	none	66	AL110152
cDNA FLJ14162 fis, clone NT2RM4002504	Increase	67	none	68	AK024224
cDNA DKFZp434E1723 (clone DKFZp434E1723)	Increase	69	none	70	AL137473
cDNA FLJ11041 fis, clone PLACE1004405	Increase	71	none	72	AK001903
cDNA FLJ10433 fis NT2RP1000478	Increase	73	none	74	AK001295
cDNA DKFZp434O071	Increase	75	none	76	AF125392
cDNA FLJ23313 fis, clone HEP11919	Increase	77	none	78	AK026966
ESTs	Increase	79	none	80	R62339
ESTs	Increase	81	none	82	AA489477
ESTs	Increase	83	none	84	R06601
ESTs	Increase	85	none	86	R00332
ESTs	Increase	87	none	88	AA463469
ESTs	Increase	89	none	90	H56028
ESTs	Increase	91	none	92	AA293300
ESTs	Increase	93	none	94	AW250104
ESTs	Increase	95	none	96	BE382614
ESTs	Increase	97	none	98	H59618
ESTs	Increase	99	none	100	AA449703
ESTs	Increase	101	none	102	AA521311
ESTs	Increase	103	none	104	W69170

ESTs	Increase	105	none	106	R51835
ESTs	Increase	107	none	108	H87770
ESTs	Increase	109	none	110	R69248
ESTs	Increase	111	none	112	T68844
ESTs	Increase	113	none	114	AA454177
ESTs	Increase	115	none	116	AA026562
ESTs	Increase	117	none	118	T73780
ESTs	Increase	119	none	120	AA401496
ESTs	Increase	121	none	122	AA489636
ESTs	Increase	123	none	124	AA446361
ESTs	Increase	125	none	126	AA931411
ESTs	Increase	127	none	128	R24223
ESTs	Increase	129	none	130	R22252
ESTs	Increase	131	none	132	AA612751
ESTs	Increase	133	none	134	AW964331
ESTs	Increase	135	none	136	AI018611
ESTs	Increase	137	none	138	AA451886
ESTs	Increase	139	none	140	R06520
ESTs	Increase	141	none	142	T48278
ESTs	Increase	143	none	144	R68736
cDNA FLJ14028 fis, clone HEMBA1003838	Repression	145	none	146	AK024090
cDNA DKFZp564D016 (clone DKFZp564D016)	Repression	147	none	148	AL050021
cDNA FLJ11302 fis, clone PLACE1009971	Repression	149	none	150	AK002164
NEDO FLJ10309 fis cl NT2RM2000287	Repression	151	none	152	AK001171
Sequence from clone RP11-39402 on ch 20	Repression	153	none	154	AK022731
ESTs	Repression	155	none	156	AA420992
ESTs	Repression	157	none	158	AA693797
ESTs	Repression	159	none	160	AA456437
ESTs	Repression	161	none	162	AA429367
ESTs	Repression	163	none	164	AA434382

ESTs	Repression	165	none	166	AA664228
ESTs	Repression	167	none	168	R44397
ESTs	Repression	169	none	170	AA923509
ESTs	Repression	171	none	172	W87747
ESTs	Repression	173	none	174	AA973568
ESTs	Repression	175	none	176	T98529
ESTs	Repression	177	none	178	AA022679
ESTs	Repression	179	none	180	H17921
ESTs	Repression	181	none	182	R00766
ESTs	Repression	183	none	184	W91958
ESTs	Repression	185	none	186	R63694
ESTs	Repression	187	none	188	AA425386
ESTs	Repression	189	none	190	AA909912
ESTs	Repression	191	none	192	T99032
ESTs	Repression	193	none	194	H52503
ESTs	Repression	195	none	196	AA127017
ESTs	Repression	197	none	198	R38647
ESTs	Repression	199	none	200	T87233
ESTs	Repression	201	none	202	AA130351
ESTs	Repression	203	none	204	H49601
ESTs	Repression	205	none	206	AA598952
ESTs	Repression	207	none	208	AA991868
ESTs	Repression	209	none	210	T60111
ESTs	Repression	211	none	212	AA897090
ESTs	Repression	213	none	214	AA679939
ESTs	Repression	215	none	216	AA630167
BCL2/adenovirus E1B 19kD-interacting protein 3-like	Increase	217	NP_004322	218	NM_004331
Solute carrier family 2, member 1	Increase	219	NP_006507	220	NM_006516
PDGF beta	Increase	221	NP_002599	222	NM_002608
lactate dehydrogenase A	Increase	223	NP_005557	224	NM_005566

Tissue factor	Increase	225	NP_001984	226	NM_001993
Vascular endothelial growth factor	Increase	227	NP_003367	228	NM_003376
RTP / NDRG1	Increase	229	NP_006087	230	NM_006096
Procollagen-proline 4-hydroxylase alpha 1	Increase	231	NP_000908	232	NM_000917
BC12/adenovirus E1B-interacting protein 3	Increase	233	NP_004043	234	NM_004052
Plasminogen activator inhibitor, type I	Increase	235	AA060003	236	M16006
Cyclooxygenase 2	Increase	237	AA057317	238	U04636
Metallothionein IH	Increase	239	CAA46046	240	X64834
Metallothionein IL	Increase	241	P80297	242	AJ011772
Metallothionein-IG	Increase	243	AA059873	244	J03910
Metallothionein IE (functional)	Increase	245	AA059587	246	M10942
Solute carrier family 2, member 3	Increase	247	AA061083	248	M20681
Hexokinase 2	Increase	249	CAA86511	250	Z46376
Interleukin 8	Increase	251	CAA68742	252	Y00787
Glyceraldehyde-3-phosphate dehydrogenase	Increase	253	NP_002037	254	NM_002046
Phosphoglycerate kinase I	Increase	255	NP_000282	256	NM_000291
Enolase I	Increase	257	NP_001419	258	NM_001428
aldolase C, fructose-bisphosphate (ALDOC)	Increase	259	NP_005156	260	NM_005165
Triosephosphate isomerase 1 (TPI1)	Increase	261	NP_000356	262	NM_000365
Adenylate kinase 3 (AK3)	Increase	263	NP_037542	264	NM_013410
Metallothionein-2a	Increase	265	AA059583	266	J00271
Osteopontin	Increase	267	CAA31984	268	X13694
Granulin	Increase	269	AA058617	270	AK000607
Hypoxia-inducible protein 2	Increase	271	NP_037464	272	NM_013332
Enolase 2, (gamma, neuronal)	Increase	273	NP_001966	274	NM_001975
Glycogen synthase I (muscle)	Increase	275	AA060385	276	U32573
Activated leucocyte cell adhesion molecule	Increase	277	NP_001618	278	NM_001627
MAX-interacting protein 1	Increase	279	NP_005953	280	NM_005962
Nuclear receptor co-repressor	Increase	281	NP_006302	282	NM_006311
Chitinase 3-like 2	Increase	283	AA050597	284	U49835
BACH1 transcription factor	Increase	285	NP_001177	286	NM_001186

Phosphoglucomutase 1	Increase	287	NP_002624	288	NM_002633
CGI-109 protein	Increase	289	AAD34104	290	AF151867
SAP30	Increase	291	NP_003855	292	NM_003864
ATP-binding cassette transporter-1	Increase	293	NP_005493	294	NM_005502
SEC24 protein	Increase	295	CAA10334	296	AJ131244
Trinucleotide repeat containing 3	Increase	297	NP_005869	298	NM_005878
Post-synaptic density protein 95	Increase	299	AAC52113	300	U83192
Tumor protein D52	Increase	301	NP_005070	302	NM_005079
Cyclin-dependent kinase inhibitor p27kip1	Increase	303	NP_004055	304	NM_004064
phosphoinositide-3-kinase, catalytic, beta	Increase	305	NP_006210	306	NM_006219
Solute carrier family 5, member 3	Increase	307	NP_008864	308	NM_006933
PSCDBP	Increase	309	NP_004279	310	NM_004288
Solute carrier family 2, member 5	Increase	311	AAA52570	312	M55531
Adipophilin	Increase	313	NP_001113	314	NM_001122
Early development regulator 2	Increase	315	NP_004418	316	NM_004427
B-cell translocation gene 1,	Increase	317	NP_001722	318	NM_001731
SH3PX1	Increase	319	NP_057308	320	NM_016224
Cyclin G2	Increase	321	NP_004345	322	NM_004354
NAG-5 protein	Increase	323	NP_057530	324	NM_016446
Cytochrome P450 IB1 (dioxin-inducible)	Increase	325	NP_000095	326	NM_000104
Butyrate response factor 1	Increase	327	NP_004917	328	NM_004926
p8 protein (candidate of metastasis 1)	Increase	329	NP_036517	330	NM_012385
chemokine (C-X-C motif), receptor 4 (CXCR4)	Increase	331	NP_003458	332	NM_003467
solute carrier family 16, member 6	Increase	333	AAC52014	334	U79745
Proline-rich protein with nuclear targeting signal (B4-2)	Increase	335	NP_006804	336	NM_006813
RNA helicase-related protein	Increase	337	AAC32396	338	AF083255
Cytochrome P450, subfamily XXVIIIB, polypeptide 1	Increase	339	BAA22656	340	AB005989
SHB adaptor protein	Increase	341	CAA53091	342	X75342
Papillomavirus regulatory factor (PRF-1)	Increase	343	NP_061130	344	NM_018660
SLC31A2/hCTR1	Increase	345	NP_001851	346	NM_001860
UDP-glucose pyrophosphorylase 2 (UGP2)	Increase	347	NP_006750	348	NM_006759

Proline 4-hydroxylase, alpha polypeptide II	Increase	349	NP_004190	350	NM_004199
Stearoyl-CoA desaturase	Increase	351	BAA93510	352	AB032261
Diacylglycerol kinase, zeta	Increase	353	NP_003637	354	NM_003646
Serine protease II	Increase	355	BAA13322	356	Y07921
IL-1 receptor antagonist, alternatively spliced forms	Increase	357	AAB92268, AAB92269, AAB92270	358	U65590
NSI-binding protein	Increase	359	NP_006460	360	NM_006469
Activin A receptor type I	Increase	361	NP_001096	362	NM_001105
FGF receptor activating protein I (FRAG1)	Increase	363	AAF19156	364	AF159621
Galectin-8	Increase	365	AAF19370	366	AF193806
Glucose 6-phosphate isomerase	Increase	367	NP_000166	368	NM_000175
D123 protein	Increase	369	AAC34738	370	U27112
Decl.	Increase	371	NP_003661	372	NM_003670
Rab-8b	Increase	373	NP_057614	374	NM_016530
BL34	Increase	375	AAB26289	376	S59049
Polyubiquitin UbC	Increase	377	BAA23632	378	AB009010
Integrin alpha 5	Increase	379	NP_002196	380	NM_002205
Jk-recombination signal binding protein	Increase	381	AA60258	382	L07872
DEAD-box protein abstrakt	Increase	383	NP_057306	384	NM_016222
High mobility group 2 protein	Increase	385	AA58659	386	M83665
Decidual protein induced by progesterone	Increase	387	NP_008952	388	NM_007021
GM2 ganglioside activator protein.	Increase	389	CAA43993, CAA43994	390	X62078
CCR4 associated factor I (CAF1)	Increase	391	AAD02685	392	AF053318
Nucleoside phosphorylase	Repression	393	NP_000261	394	NM_000270
Monocyte chemotactic protein 1	Repression	395	NP_002973	396	NM_002982
Endothelin I	Repression	397	NP_001946	398	NM_001955
Heat shock 70kD protein 4	Repression	399	AA02807	400	L12723
Annexin AI	Repression	401	NP_000691	402	NM_000700

p67 myc protein	Repression	403	CAA25105, CAA25106	404	X00364
Alpha-2-macroglobulin	Repression	405	NP_000005	406	NM_000014
Macrophage inflammatory protein 1b	Repression	407	NP_002975	408	NM_002984
Sex hormone-binding globulin	Repression	409	NP_001031	410	NM_001040
ATP-binding cassette, sub-family E (OABP), member 1	Repression	411	NP_002931	412	NM_002940
Chaperonin / Tcp zeta 1	Repression	413	NP_001753	414	NM_001762
Colony stimulating factor 1 (macrophage)	Repression	415	AAA59573	416	M27087
Dendritic cell protein (GA17)	Repression	417	NP_006351	418	NM_006360
G protein-coupled receptor 44	Repression	419	NP_004769	420	NM_004778
Keratin 6A	Repression	421	NP_005545	422	NM_005554
lymphocyte adaptor protein	Repression	423	NP_005466	424	NM_005475
Neuro-oncological ventral antigen 1	Repression	425	AAA16022	426	U04840
N-SMase / FAN	Repression	427	CAA65405	428	X96586
Peptidylprolyl isomerase F (cyclophilin F)	Repression	429	NP_005720	430	NM_005729
PLECKSTRIN	Repression	431	NP_002655	432	NM_002664
High affinity immunoglobulin epsilon receptor beta	Repression	433	AAF17243	434	AF201951
Ribosomal protein L44	Repression	435	NP_000992	436	NM_001001
Solute carrier family 6 No1	Repression	437	NP_003033	438	NM_003042
Synaptopodin	Repression	439	NP_009217	440	NM_007286
TERA protein	Repression	441	AAF87322	442	AF212220
TGF beta-stimulated protein TSC-22	Repression	443	NP_006013	444	NM_006022
Tubulin, beta, 2	Repression	445	NP_006079	446	NM_006088
Calgranulin A	Repression	447	NP_002955	448	NM_002964
Replication factor C (145 KDa)	Repression	449	NP_002904	450	NM_002913
Signal recognition particle 19 kD protein	Repression	451	NP_003126	452	NM_003135
Transcription factor SUPT3H	Repression	453	NP_003590	454	NM_003599
Proteasome component C9	Repression	455	NP_002780	456	NM_002789
Maf-related leucine zipper homolog	Repression	457	NP_005452	458	NM_005461
dynein, cytoplasmic, light intermediate polypeptide 2	Repression	459	NP_006132	460	NM_006141
Heterochromatin-like protein 1	Repression	461	NP_057671	462	NM_016587

Monocyte chemotactic protein 3	Repression	463	NP_006264	464	NM_006273
Fatty-acid-Coenzyme A ligase, long-chain 2	Repression	465	BA A00931	466	D10040
Programmed cell death 5 / TFAR19	Repression	467	NP_004699	468	NM_004708
Small inducible cytokine A3	Repression	469	AA A36316	470	M23452
Cytochrome c oxidase subunit VIc	Repression	471	NP_004365	472	NM_004374
NASP histone-binding prot.	Repression	473	NP_002473	474	NM_002482
Ecotropic viral integration site 2A	Repression	475	NP_055025	476	NM_014210
Sjogren syndrome antigen B	Repression	477	AA A51885	478	J04205
Monocyte chemotactic protein 2	Repression	479	NP_005614	480	NM_005623
GRO2/ macrophage inflammatory protein 2a	Repression	481	NP_002080	482	NM_002089
Small nuclear ribonucleoprotein SM D1	Repression	483	NP_008869	484	NM_006938
GRO1/ macrophage inflammatory protein 2 precursor	Repression	485	NP_001502	486	NM_001511
Lymphocyte adhesion molecule 1	Repression	487	NP_000646	488	NM_000655

TABLE 27 cross-references all protein and nucleotide sequences (SEQ ID Nos.) that are referenced herein to accession numbers in public databases available as of 8.12.01.

Clone ID	New Name	Old Name	Protein Seq ID	Protein Accession	Nucleotide Seq ID	GenBank Locus
p1F12	Hypothetical protein FLJ13611	cDNA FLJ13611 fis, clone PLACE1010802	1	NP_079217	2	NM_024941
p1F2	Hypothetical protein FLJ20037	Hypothetical protein FLJ20037	3	CAB65981	4	NM_017633
p1F10	Hypothetical protein DKFZp434P0116	hypothetical protein DKFZp434P0116	5	T46364	6	NM_017593
p1F19	Hypothetical protein KIAA0212	KIAA0212	7	BAA13203	8	NM_014674
p1F8	Hypothetical protein KIAA0914	KIAA0914	9	NP_055698	10	NM_014883
p1F5	Hypothetical protein FLJ20281	Hypothetical protein FLJ20281	11	XP_008736	12	NM_017742
p1F18	Hypothetical protein KIAA0876	KIAA0876	13	BAA74899	14	XM_035625
p1F7	Spectrin, beta, non-erythrocytic 1	cDNA FLJ13700 fis, clone PLACE2000216	15	NP_003119	16	NM_003128
p1F21	Hematopoietic Zinc finger protein	DKFZP586G1122 protein	17	AAL08625	18	AK024404
p1F9	Hypothetical protein KIAA0742	Putative zinc finger protein LOC55818	19	NP_060903	20	AB018285
p1E13	Hypothetical protein PRO0823	hypothetical protein PRO0823	21	AAF71073	22	AF116653
p1D1	Hypothetical protein FLJ10134	Hypothetical protein FLJ10134	23	NP_060474	24	NM_018004
p1D2	Hypothetical protein FLJ10134	Hypothetical protein FLJ10134	23	NP_060474	24	NM_018004
p1D4	Hypothetical protein FLJ20500	Hypothetical protein FLJ20500	25	NP_061931	26	NM_019058
p1D9	Hypothetical protein DKFZP564D116	DKFZP564D116 protein	27	T08708	28	AL050022
p1D12	Hypothetical protein KIAA1376	KIAA1376 protein	29	BAA92614	30	AB037797
p1D15	TRIP-Br2	Hypothetical protein KIAA0127	31	NP_055570	32	NM_014755
p1D16	Hypothetical protein FLJ20308	Hypothetical protein FLJ20308	33	XP_039552	34	AK000315
p1J13	Hypothetical nuclear factor SBB122	Hypothetical nuclear factor SBB122	35	NP_065128	36	NM_020395
p1J22	Hypothetical protein KIAA1429	DKFZP434J116 protein	37	BAA92667	38	AB037850
p1I16	Hypothetical protein FLJ10206	Hypothetical prot. FLJ10206	39	AAH06108	40	NM_018025

p115	Hypothetical protein FLJ10815	hypothetical protein FLJ10815	41	BAA91830	42	NM_018231
p113	Hypothetical protein FLJ11100	Hypothetical protein FLJ11100	43	NP_060701	44	NM_018321
p117	Hypothetical protein FLJ20644	hypothetical protein FLJ2064	45	NP_060387	46	NM_017917
p115	Hypothetical protein CGI-117	Hypothetical protein HSPC111	47	Q9Y3C1	48	NM_016391
p117	Uridine 5' monophosphate hydrolase 1	hypothetical protein LOC51251	49	NP_057573	50	NM_016489
	Hypothetical protein KIAA0014	KIAA0014	51	NP_055480	52	NM_014665
p114	Hypothetical protein HSPC196	Hypothetical protein HSPC196	53	NP_057548	54	NM_016464
p118	Hypothetical protein FLJ11296	Hypothetical protein FLJ11296	55	XP_004747	56	NM_018384
p116	Hypothetical protein KIAA1668	Hypothetical protein bA39SL14	57	BAB33338	58	AB051455
p111	SECIS binding protein 2	cDNA FLJ13016 fis, clone NT2RP3000624	59	AAK57518	60	AF380995
p1E8	cDNA: FLJ22249 fis, clone HRC02674	cDNA DKFZp586H0324 clone DKFZp586H0324	61	None	62	AK025902
p1E8	Plexin C1	Clone 23785	63	NP_005752	64	NM_005761
p1E16	cDNA DKFZp586E1624	cDNA DKFZp586E1624	65	None	66	AL110152
p1D5	ERO1 (S. cerevisiae)-like	cDNA FLJ14162 fis, clone NT2RM4002504	67	NP_055399	68	NM_014584
p1D6	ERO1 (S. cerevisiae)-like	cDNA FLJ14162 fis, clone NT2RM4002504	67	NP_055399	68	NM_014584
p1E12	Hypothetical protein DKFZp434E1723	cDNA DKFZp434E1723 (clone DKFZp434E1723)	69	XP_05338	70	BC010005
p1E10	cDNA FLJ11041 fis, clone PLACE1004405	cDNA FLJ11041 fis, clone PLACE1004405	71	None	72	AK001903
p1C21	Tubulin, beta, 4	cDNA FLJ10433 fis, clone NT2RP1000478	73	NP_006077	74	NM_006086
p1D10	Insulin induced protein 2	cDNA DKFZp434O071	75	AAD43048	76	AF125392
p1D13	Adenylate kinase 3	cDNA FLJ23313 fis, clone HEP11919	77	NP_037542	78	NM_013410
p1E9	Novel PI-3-kinase adapter	ESTs	79	None	80	R62339
p1F1	EST	ESTs	81	None	82	AA489477
p1E7	Novel Metallothionein	ESTs	83	None	84	R06601

p1E6	EGL nine (C.elegans) homolog 3	ESTs	85	NP_071356	86	NM_022073
p2B1	PRAME	ESTs	87	NP_006106	88	NM_006115
p1D14	Clorf12	ESTs	89	NP_071334	90	NM_022051
p1D17	Semaphorin 4b	ESTs	91	BAB21836	92	AB051532
p1P14	Semaphorin 4b	ESTs	91	BAB21836	92	AB051532
p1C24	SLC25A19	ESTs	93	NP_068380	94	NM_021734
p1D3	Serine carboxypeptidase 1	ESTs	95	NP_067639	96	NM_021626
p1E14	Unknown mRNA (schizophrenia-linked)	ESTs	97	None	98	AY010112
p1E20	Myo-inositol monophosphatase A3	ESTs	99	AAK52336	100	NM_017813
p2A24	EST	ESTs	101	None	102	AA521314
p1E17	Hypothetical protein FLJ1668	ESTs	103	BAB71124	104	AK056230
p1E19	EST	ESTs	105	None	106	R51835
p1E15	cDNA Y127F12	ESTs	107	None	108	AF075018
p1E11	EST	ESTs	109	None	110	R69248
p1E23	cDNA FLJ14041 fis, clone HEMBA1005780	ESTs	111	None	112	AK024103
p1E21	Glutamate-cysteine ligase, modifier subunit	ESTs	113	NP_002052	114	NM_002061
p1D23	PTEN	ESTs	115	NP_000305	116	NM_000314
p1D24	EST	ESTs	117	None	118	T73780
p1D22	MAX-interacting protein 1	ESTs	119	NP_005953	120	NM_005962
p1E2	Mannosidase, alpha, class 1A, member 1	ESTs	121	NP_005898	122	NM_005907
p1E1	EST	ESTs	123	None	124	AA446361
p1E4	EST	ESTs	125	None	126	AA931411
p1D18	cDNA FLJ13443 fis, clone PLACE1002853	ESTs	127	None	128	AK023505
p1D21	Hypothetical protein FLJ22622	ESTs	129	BAB15424	130	NM_025151
p1C22	CD84-H1	ESTs	131	AAK69052	132	AF275725
p1C23	Hypothetical protein FLJ12832	ESTs	133	XP_043394	134	AK022894
p1D11	EST	ESTs	135	None	136	AA251748

pIE3	CYP1B1	ESTs	137	NP_000095	138	NM_000104
pID20	Hypothetical protein KIAA1125	ESTs	139	XP_012932	140	AB032951
pIE5	Hepcidin antimicrobial peptide	ESTs	141	NP_066998	142	NM_021175
pID19	EST	ESTs	143	None	144	R68736
p2A15	Sialyltransferase	cDNA FLJ14028 fis, clone HEMBA1003838	145	NP_006447	146	NM_006456
p1I14	cDNA DKFZp564D016	cDNA DKFZp564D016 (clone DKFZp564D016)	147	None	148	AL050021
p1I2	cDNA FLJ11302 fis, clone PLACE1009971	cDNA FLJ11302 fis, clone PLACE1009971	149	None	150	AK002164
p1I12	Hypothetical protein MGC4549	NEDO FLJ10309 fis cl NT2RM2000287	151	XP_032794	152	NM_032377
p13	ELMO2	Sequence from clone RP11-39402 on ch 20	153	AAL14467	154	XM_012933
p1I10	EST	ESTs	155	None	156	AA420992
p1H18	Ubiquitin specific protease 7	ESTs	157	NP_003461	158	NM_003470
p1H24	Nucleolar phosphoprotein Nopp34	ESTs	159	NP_115766	160	NM_032390
p1E22	cDNA FLJ13618 fis, clone PLACE1010925	ESTs	161	None	162	AK023680
p1H21	Hypothetical protein FLJ13511	ESTs	163	NP_149014	164	NM_033025
p1I1	Ribosomal RNA intergenic spacer	ESTs	165	None	166	AA664228
p1H14	EST	ESTs	167	None	168	R44397
p1H11	Carboxypeptidase M	ESTs	169	NP_001865	170	NM_001874
p1H17	EST	ESTs	171	None	172	W87747
p1H12	EST	ESTs	173	None	174	AA973568
p1H7	EST	ESTs	175	None	176	T98529
p1H15	EST	ESTs	177	None	178	AA022679
p1H20	EST	ESTs	179	None	180	H17921
p1H8	ABL	ESTs	181	NP_009297	182	NM_007313
p1H16	EST	ESTs	183	None	184	W91958
p1H9	EST	ESTs	185	None	186	R63694

p1H23	Hypothetical protein FLJ21094	ESTs	187	AAH14003	188	AK024747
p1H10	EST	ESTs	189	None	190	AA909912
p1H6	EST	ESTs	191	None	192	T99032
p1H13	EST	ESTs	193	None	194	H52503
p1H19	EST	ESTs	195	None	196	AA127017
p1G22	EST	ESTs	197	None	198	R38647
p1G21	EST	ESTs	199	None	200	T87233
p1H1	Hypothetical protein FLJ10826	ESTs	201	BAB14226	202	NM_018233
p1G20	cDNA YO23H03	ESTs	203	None	204	AF075053
p1H5	Hypothetical protein FLJ22690	ESTs	205	NP_078987	206	NM_024711
p1G19	Mitochondrion sequence	ESTs	207	AAH05845	208	BC005845
p1H2	Fatty acid binding protein 5	ESTs	209	NP_001435	210	NM_001444
p1G18	Mitochondrion sequence	ESTs	211	None	212	BC001612
p1H4	EST	ESTs	213	None	214	AA679939
p1H3	EST	ESTs	215	None	216	AA630167
	BCL2/adenovirus E1B 19kD-interacting protein 3-like	BCL2/adenovirus E1B 19kD-interacting protein 3-like	217	NP_004322	218	NM_004331
	SLC2A1	Solute carrier family 2, member 1	219	NP_006507	220	NM_006516
p1P3	PDGFB	PDGF beta	221	NP_148937	222	NM_033016
p1A8	Lactate dehydrogenase A	lactate dehydrogenase A	223	NP_005557	224	NM_005566
p1A9	Lactate dehydrogenase A	lactate dehydrogenase A	223	NP_005557	224	NM_005566
p1B17	Tissue factor	Tissue factor	225	NP_001984	226	NM_001993
p1O20	VEGF	Vascular endothelial growth factor	227	NP_003367	228	NM_003376
p1B2	N-myc downstream regulated	RTP / NDRG1	229	NP_006087	230	NM_006096
p1B3	Proline 4-hydroxylase, alpha polypeptide 1	Procollagen-proline 4-hydroxylase alpha 1	231	NP_000908	232	NM_000917
	BCL2/adenovirus E1B-interacting protein 3	BCL2/adenovirus E1B-interacting protein 3	233	NP_004043	234	NM_004052
p1B18	Plasminogen activator inhibitor, type 1	Plasminogen activator inhibitor, type 1	235	NP_000593	236	NM_000602
p1B19	Plasminogen activator inhibitor, type 1	Plasminogen activator inhibitor, type 1	235	NP_000593	236	NM_000602

	1	type I				
pIN17	COX-2	Cyclooxygenase 2	237	NP_000954	238	NM_000963
pIA24	Metallothionein 1H	Metallothionein 1H	239	NP_005942	240	NM_005951
	Metallothionein 1L	Metallothionein 1L	241	NP_002441	242	NM_002450
pIB1	Metallothionein 1G	Metallothionein-1G	243	NP_005941	244	NM_005950
	Metallothionein 1E (functional)	Metallothionein 1E (functional)	245	None	246	AA872383
pIA1	SLC2A3	Solute carrier family 2, member 3	247	NP_008862	248	NM_006931
pIA2	SLC2A3	Solute carrier family 2, member 3	247	NP_008862	248	NM_006931
pIA3	SLC2A3	Solute carrier family 2, member 3	247	NP_008862	248	NM_006931
pIA4	SLC2A3	Solute carrier family 2, member 3	247	NP_008862	248	NM_006931
pIA15	Hexokinase-2	Hexokinase 2	249	NP_000180	250	NM_000189
pIA16	Hexokinase-2	Hexokinase 2	249	NP_000180	250	NM_000189
pIA17	Hexokinase-2	Hexokinase 2	249	NP_000180	250	NM_000189
pIA18	Hexokinase-2	Hexokinase 2	249	NP_000180	250	NM_000189
pIB14	Interleukin 8	Interleukin 8	251	NP_000375	252	NM_000584
pIB15	Interleukin 8	Interleukin 8	251	NP_000375	252	NM_000584
pIB16	Interleukin 8	Interleukin 8	251	NP_000375	252	NM_000584
pIA11	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	253	NP_002037	254	NM_002046
pIA12	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	253	NP_002037	254	NM_002046
pIA13	Phosphoglycerate kinase 1	Phosphoglycerate kinase 1	255	NP_000282	256	NM_000291
pIA14	Enolase I	Enolase I	257	NP_001419	258	NM_001428
pIA19	Aldolase C	aldolase C, fructose-bisphosphate (ALDOC)	259	NP_005156	260	NM_005165
pIA20	Triosephosphate isomerase 1	Triosephosphate isomerase 1 (TPI1)	261	NP_000356	262	NM_000365
pIA22	Adenylate kinase 3	Adenylate kinase 3 (AK3)	263	NP_037542	264	NM_013410
pIA23	Metallothionein 2A	Metallothionein-2a	265	NP_005944	266	NM_005953
pIB20	Osteopontin	Osteopontin	267	NP_000573	268	NM_000582
pIB21	Osteopontin	Osteopontin	267	NP_000573	268	NM_000582
pIC17	Granulin	Granulin	269	NP_002078	270	NM_002087
pIC18	Granulin	Granulin	269	NP_002078	270	NM_002087

p1D8	Hypoxia-inducible protein 2	Hypoxia-inducible protein 2	271	NP_037464	272	NM_013332
p1A10	Enolase 2	Enolase 2, (gamma, neuronal)	273	NP_001966	274	NM_001975
p1G24	Glycogen synthase 1	Glycogen synthase 1 (muscle)	275	NP_002094	276	NM_002103
p1G23	ALCAM	Activated leucocyte cell adhesion molecule	277	NP_001618	278	NM_001627
p1G5	MAX-interacting protein 1	MAX-interacting protein 1	279	NP_005953	280	NM_005962
p1G7	EST	Nuclear receptor co-repressor	281	None	282	BC008022
p2A23	Chitinase 3-like 2	Chitinase 3-like 2	283	NP_003991	284	NM_004000
p1G1	BACH1	BACH1 transcription factor	285	NP_001177	286	NM_001186
p1G15	Phosphoglucosylase 1	Phosphoglucosylase 1	287	NP_002624	288	NM_002633
p1F23	Hypothetical protein LOC51014	CGI-109 protein	289	Q9Y3B3	290	AF151867
p1G8	Sin3-associated polypeptide	SAP30	291	NP_003855	292	NM_003864
p1G13	ABCA1	ATP-binding cassette transporter-1	293	NP_003493	294	NM_003502
p1G10	SEC24 member A	SEC24 protein	295	CAA10334	296	AJ131244
p1F24	Glia-derived nexin	Trinucleotide repeat containing 3	297	AAA35883	298	M17783
p1G2	Postsynaptic density-95	Post-synaptic density protein 95	299	NP_001356	300	NM_001365
p1G11	Tumor protein D52	Tumor protein D52	301	NP_003070	302	NM_003079
p1G16	p27, Kip1	Cyclin-dependent kinase inhibitor p27kip1	303	NP_004055	304	NM_004064
p1G9	PI-3-kinase, catalytic, beta polypeptide	phosphoinositide-3-kinase, catalytic, beta	305	NP_006210	306	NM_006219
p1G4	SLC5A3	Solute carrier family 5, member 3	307	AAC39548	308	AF027153
p1G14	Cytohesin binding protein	PSCDBP	309	NP_004279	310	NM_004288
p1A5	SLC2A5	Solute carrier family 2, member 5	311	NP_003030	312	NM_003039
p1A6	SLC2A5	Solute carrier family 2, member 5	311	NP_003030	312	NM_003039
p1B6	Adipophilin	Adipophilin	313	NP_001113	314	NM_001122
p1B7	Adipophilin	Adipophilin	313	NP_001113	314	NM_001122
p1B8	Adipophilin	Adipophilin	313	NP_001113	314	NM_001122
p1B9	Adipophilin	Adipophilin	313	NP_001113	314	NM_001122
p1G17	Early development regulator 2	Early development regulator 2	315	NP_004418	316	NM_004427
p1G3	B-cell translocation gene 1	B-cell translocation gene 1,	317	NP_001722	318	NM_001731
p1F22	Sorting nexin 9	SH3PX1	319	NP_057308	320	NM_016224

p1G12	Cyclin G2	Cyclin G2	321	NP_004345	322	NM_004354
p1F11	Hypothetical protein LOC51754	NAG-5 protein	323	XP_049657	324	AL137430
p1F16	CYP1B1	Cytochrome P450 1B1 (dioxin-inducible)	325	NP_000095	326	NM_000104
p1F14	Butyrate response factor 1	Butyrate response factor 1	327	NP_004917	328	NM_004926
p1F17	P8 protein (candidate of metastasis 1)	p8 protein (candidate of metastasis 1)	329	NP_036517	330	NM_012385
p1C1	CXCR4	chemokine (C-X-C motif), receptor 4 (CXCR4)	331	NP_003458	332	NM_003467
p1C2	CXCR4	chemokine (C-X-C motif), receptor 4 (CXCR4)	331	NP_003458	332	NM_003467
p1F3	Hypothetical protein XP_017131	solute carrier family 16, member 6	333	XP_017131	334	XM_017131
p1F20	Proline-rich protein with nuclear targeting signal	Proline-rich protein with nuclear targeting signal (B4-2)	335	NP_006804	336	NM_006813
p1F6	Hypothetical protein hqp0376	RNA helicase-related protein	337	T08745	338	AF078844
p1F4	CYP1	Cytochrome P450, subfamily XXVIIIB, polypeptide 1	339	NP_000776	340	NM_000785
p1F15	SHB adaptor protein	SHB adaptor protein	341	NP_003019	342	NM_003028
p1F13	Papillomavirus regulatory factor PRF-1	Papillomavirus regulatory factor (PRF-1)	343	NP_061130	344	AK023418
p1A7	SLC31A2	SLC31A2/hCTR1	345	NP_001851	346	NM_001860
p1A21	UDP-glucose pyrophosphorylase 2	UDP-glucose pyrophosphorylase 2 (UGP2)	347	NP_006750	348	NM_006759
p1B4	Proline 4-hydroxylase, alpha polypeptide II	Proline 4-hydroxylase, alpha polypeptide II	349	NP_004190	350	NM_004199
p1B5	Proline 4-hydroxylase, alpha polypeptide II		349	NP_004190	350	NM_004199
p1B10	Stearoyl-CoA desaturase	Stearoyl-CoA desaturase	351	NP_005054	352	NM_005063
p1B11	Stearoyl-CoA desaturase	Stearoyl-CoA desaturase	351	NP_005054	352	NM_005063
p1B12	Stearoyl-CoA desaturase	Stearoyl-CoA desaturase	351	NP_005054	352	NM_005063
p1B13	Diacylglycerol kinase, zeta	Diacylglycerol kinase, zeta	353	NP_003637	354	NM_003646
p1B22	Protease, serine, 11	Serine protease 11	355	NP_002766	356	NM_002775

p1B23	Interleukin 1 receptor antagonist	IL-1 receptor antagonist, alternatively spliced forms	357	NP_000568	358	NM_000577
p1B24	NS1-binding protein	NS1-binding protein	359	NP_006460	360	NM_006469
p1C3	Activin A receptor, type I	Activin A receptor type I	361	NP_001096	362	NM_001105
p1C4	FGF receptor activating protein 1	FGF receptor activating protein 1 (FRAG1)	363	NP_055304	364	NM_014489
p1C5	Galectin 8	Galectin-8	365	NP_006490	366	NM_006499
p1C6	Glucose phosphate isomerase	Glucose 6-phosphate isomerase	367	NP_000166	368	NM_000175
p1C7	D123	D123 protein	369	NP_006014	370	NM_006023
p1C8	DEC-1	Dec1	371	NP_003661	372	NM_003670
p1C9	RAB-8b protein	Rab-8b	373	NP_057614	374	NM_016530
p1C10	Regulator of G-protein signalling 1	BL34	375	NP_002913	376	NM_002922
p1C11	Polyubiquitin	Polyubiquitin UbC	377	BAA23632	378	AB009010
p1C12	Integrin, alpha 5	Integrin alpha 5	379	NP_002196	380	NM_002205
p1C13	Jk-recombination signal binding protein	Jk-recombination signal binding protein	381	AAA60258	382	L07872
p1C14	Abstrakt	DEAD-box protein abstrakt	383	NP_057306	384	NM_016222
p1C15	High-mobility group protein 2	High mobility group 2 protein	385	NP_002120	386	NM_002129
p1C16	Decidual protein induced by progesterone	Decidual protein induced by progesterone	387	NP_008952	388	NM_007021
p1C19	GM2 ganglioside activator protein	GM2 ganglioside activator protein.	389	NP_000396	390	NM_000405
p1C20	CNOT8	CCR4 associated factor 1 (CAF1)	391	NP_004770	392	NM_004779
p1P5	Similar to Nucleoside phosphorylase	Nucleoside phosphorylase	393	None	394	AA430382
p2L23	SCYA2	Monocyte chemotactic protein 1	395	NP_002973	396	NM_002982
p1K9	Endothelin 1	Endothelin 1	397	NP_001946	398	NM_001955
p1K23	Similar to Heat shock 70kD protein 4	Heat shock 70kD protein 4	399	None	400	AA633656
p1K15	Lipocortin I	Annexin A1	401	NP_000691	402	NM_000700
p1K8	MYC	p67 myc protein	403	NP_002458	404	NM_002467
	Alpha-2-macroglobulin	Alpha-2-macroglobulin	405	NP_000005	406	NM_000014
	SCYA4	Macrophage inflammatory protein	407	XP_008449	408	XM_008449

		lb					
p1M24	Sex hormone-binding globulin	Sex hormone-binding globulin					
p1K7	ATP-binding cassette E1	ATP-binding cassette, sub-family E (OABP), member 1	409	NP_001031	410	NM_001040	
p1K16	CCT6A	Chaperonin / Top zeta 1	411	NP_002931	412	NM_002940	
p1K18	Colony-stimulating factor 1	Colony stimulating factor 1 (macrophage)	413	NP_001753	414	NM_001762	
p1N1	GA17	Dendritic cell protein (GA17)	415	AA52117	416	M37435	
p1K22	GPR44	G protein-coupled receptor 44	417	NP_006351	418	NM_006360	
p1K14	Keratin 6B	Keratin 6A	419	NP_004769	420	NM_004778	
p1K13	Lymphocyte adaptor protein	Lymphocyte adaptor protein	421	NP_005546	422	NM_005555	
p1J20	Neuro-oncological ventral antigen 1	Neuro-oncological ventral antigen 1	423	NP_005466	424	NM_005475	
p1J22	Neutral sphingomyelinase (N-SMase) activation associated factor	N-SMase / FAN	425	NP_002506	426	NM_002515	
p1K1	Cyclophilin F	Peptidylprolyl isomerase F (cyclophilin F)	427	NP_003571	428	NM_003580	
p1K3	Pleckstrin	Pleckstrin	429	NP_005720	430	NM_005729	
p1J19	CFFM4	High affinity immunoglobulin epsilon receptor beta	431	NP_002655	432	NM_002664	
p1K2	CFFM4	High affinity immunoglobulin epsilon receptor beta	433	NP_067024	434	NM_021201	
p1K5	Ribosomal protein L36a	Ribosomal protein L44	433	NP_067024	434	NM_021201	
p1J17	SLC6A1	Solute carrier family 6 No1	435	NP_000992	436	NM_001001	
p1J18	Synaptopodin	Synaptopodin	437	NP_003033	438	NM_003042	
p1J15	TERA protein	TERA protein	439	NP_009217	440	NM_007286	
p1K4	TSC-22	TGF beta-stimulated protein TSC-22	441	NP_067061	442	NM_021238	
p2A14	EST	Tubulin, beta, 2	443	NP_006013	444	NM_006022	
p1J23	Calgranulin A	Calgranulin A	445	None	446	AA988110	
p1J21	Replication factor C large subunit	Replication factor C (145 KDa)	447	NP_002955	448	NM_002964	
p1J24	Signal recognition particle 19kD	Signal recognition particle 19 kD protein	449	NP_002904	450	NM_002913	
			451	NP_003126	452	NM_003135	

p1J16	cDNA: FLJ23019 fis, clone LNG00916	Transcription factor SUPT3H	453	None	454	AK026672
p1J2	Proteasome subunit, alpha type, 4	Proteasome component C9	455	NP_002780	456	NM_002789
p1J9	MAFB	Maf-related leucine zipper homolog	457	NP_005452	458	NM_005461
p1J10	DNCLJ2	dynein, cytoplasmic, light intermediate polypeptide 2	459	NP_006132	460	NM_006141
p1J1	Chromobox homolog 3	Heterochromatin-like protein 1	461	NP_057671	462	NM_016587
p1J5	SCYA7	Monocyte chemotactic protein 3	463	NP_006264	464	NM_006273
p1J11	Fatty-acid-Coenzyme A ligase, long- chain 2	Fatty-acid-Coenzyme A ligase, long-chain 2	465	NP_066945	466	NM_021122
p1J8	Programmed cell death 5	Programmed cell death 5 / TFAR19	467	NP_004699	468	NM_004708
p1J20	SCYA3L	Small inducible cytokine A3	469	CAA36397	470	X52149
p1J3	Furin	Cytochrome c oxidase subunit VIc	471	NP_002560	472	NM_002569
p1J12	Nuclear autoantigenic sperm protein	NASP histone-binding prot.	473	NP_002473	474	NM_002482
p1J23	Ecotropic viral integration site 2A	Ecotropic viral integration site 2A	475	NP_035025	476	NM_014210
p1J7	Sjogren syndrome antigen B	Sjogren syndrome antigen B	477	NP_003133	478	NM_003142
p1J21	SCYA8	Monocyte chemotactic protein 2	479	NP_005614	480	NM_005623
p1J19	GRO2	GRO2/ macrophage inflammatory protein 2a	481	NP_002080	482	NM_002089
p1J4	Small nuclear ribonucleoprotein D1	Small nuclear ribonucleoprotein SM D 1	483	NP_008869	484	NM_006938
p1J24	GRO1	GRO1/ macrophage inflammatory protein 2 precursor	485	NP_001502	486	NM_001511
p1J18	Selectin L	Lymphocyte adhesion molecule 1	487	NP_000646	488	NM_000655

Sequence listing

1

5 MLTLPQNFNGNIFLGETFSSYISVHNSNQVVKDILVKADLQTSQRLNLSASNAVAELKPDCCIDDVIHHEVKEIGTHILVCA
VSYYTQAGEKMYFRKFFKQVLKPLDVKTKFYNAESDLSSVTDEVFLEAQIQNMTTSPMFMEKVSLEPSIMYNVTELNVSQAG
ECVSTFGSRAYLQPMQDTRQYLYCLKPKNEFAEKAGIIKGVTVIGKLDIVWKTNLGERGRLQTSQQLQRMAPGYGDVRLSLEAIPD
TVNLEEPFHITCKITNCERTMDLVLEMCNTNSIHWCGISGRQLGKLPSSSLCLALTLSSVQGLQSI SGLRLTDTFLKRTYE
YDDIAQVCVVSIAIKVES

2

10 AAAAAGTGCCGGTCAAAATGGAAGTGAATCCCCCTAAACAGGAGCACCTGCTGGCGCTAAAAGTGATGCGGCTGACTAAGCCTA
CTTTATTCACCAATATCCCAGTAACATGTGAAGAGAAAGACTTACCTGGAGATCTCTTTAACAGCTGATGAGAGATGATCCTT
CAACCGTTAATGGTGCAGAAAGTTTAAATGTTGGGATAAATGCTGACTTTACCACAGAATTTTGGGAATATATTTTGGGAGAGA
CCTTTTCCAGTTATATCAGCGTTCATAATGATAGCAATCAAGTTGTAAAAGACATATTAGTAAAAGCTGATCTTCAGACAAGTT
CTCAGCGTTTAAATCTTTAGCCCTCCAAATGCTGCAGTGGCTGAACTTAAACCGGATTGTGTATTGATGATGTCATACATCATG
15 AAGTCAAGAAATTTGGAACACACATCTTGGTATGTGCTGTGAGTTATACAACCTCAGGCTGGAGAAAAATGTATTTTCAAAAAAT
TCTTCAAAATTTCAAGTTCTCAAAACCATTTGGATGTGAAAAACCAATTTTACAATGCAGAGAGTGACCTCAGTTCTGTGACTGATG
AAGTATTTCTGGAAGCCCAGATTCAGAAATATGACAACCTCACCTATGTTTATGGAGAAGGTTTCACTGGAGCCATCTATTATGT
ACAATGTAAACAGAAATTAATTCAGTCAGCCAAGCTGGAGAATGTGTGTCTACGTTTGGGTCAAGAGCATATTTGCAACCAATGG
ATACACGCCAGTACTTATACTGCCCTAAAGCCAAAGAATGAATTTGCAGAAAAAGCAGGCATCATTAAGGAGTAACAGTAATTG
20 GAAAATTTGATATAGTATGAAAAACAAATCTAGGTGAAAGGGGAAGGTTACAGACCAGCCAACTTCAAAGAATGGCTCCAGGTT
ATGGAGATGTTAGGTTGTCTTTGGAGGCAATACCAGATACCGTAAACCTTGAAGAACCTTTTCATATTACCTGTAAAATAACAA
ACTGCAGTGAAGGAGCATGAGATCTGGTTTGGAAATGTGCAATACCAATTCATCCACTGGTGTGGAATTTCAAGGAAGACAGC
TGGGAAAAGCTGCATCCAAAGTTCTTCGCTCTGCTCTGCCCTTACTCTGCTTTCTTCAGTACAGGAGTGCAGGAGTCTCTGGCT
TAAGACTAACAGACACATCTTAAAGAGAACATATGAATATGATGACATCGCACAAAGTCTGTGTGGTATCTTCTGCCATTTAAAG
25 TGGAAAGCTGAAGGAACTTCCAATGTTAGGCTTTTCATTTAGTTTACAGAACTGCTCTTTTTGTGTACCTTTGTAAAATGATG
ACGTCAACAACGAAGTAAATATGTTACTTAAATTTTCTTTCCTGTAAAATGGTTAAAATATTAATATTTGTTTGAAGAGATC
TGATTTTATCTGTGAATTTATATTTGAAATGAACATGTGTATATTTCTACACCTATTATTTAATTTTCAATTTTATAGATGA
CCATTGGACTTTGTTCTCCAAAAGCTGTGTATCTGAGACCATTGTCCCTAGCAAGTTATCTAGAACACGAGTCAGCACACTTT
TTATGTAAGGTACAAGATACTAAATATTTTAGGCTTTGCAGGCCGTAAGGTTTCTGTACAAATACTGAACCTGCCATCGTAC
30 TTCAAGAGCAGCCAGTAGGCAAAAAAGTACAGTGGTGTACAAAAAAGAGCAGTTGGAAGTCTGGAAGCTGAGGCAGGAGAAT
CACTTGAACCGGGGAGGCAGAGGTTGCAGTTAAGCCGAGATCATGCTACTGCACCTCTAGCCTGAGCGACAGAGTGAGACTCTGT
CTCAAAAAAAGAGGCAGCTGGGCAGATTTGGCCTGCAGGGTATAGTTGGCTGACCTCAATCTAGAAGATAGAGTTGTTTCT
TAGAACTTTGTAATATTTTGGTTAGAAATATTTTCAAGTTAACTCCAGTTTTCCTAGCTATTCCGGTCAGCAGAAAAAGTTGCT
TAGAACAAAGGGGGTAGTCTCTTATAACCTGATATTTGTCAACACAGTATAAGCTTTCTAGTTGTTTCAAATTTACTGTAAT
35 GGTGTGTGCTGTATTTGTATATTTATACCTACAAACATACATATCTTTCAGCTACTGAACACATCAGAAATGGTTTACTTTTT
AGGTTGTGATTTTTTTTAAATTAACATTAAGTAAGTAATTAATTTTCTAATTATAAGCCAAAAACCCCTGGGTTTTTTTT
CTCATGATTAAGCCACTTATTTTTATAAGGTATATTTAAATAATTCCCATCTTCTGATATGATTTTAAACTCTAAATCATGTG
TTACTGTTACAAAACCTTCTCTCCATGTAATCACACTTAGTTATGAGCAAAGCAGTGAGAAAGTTGAGGTGATAAACAACCTT
TGCAGCATTAATGTAACAAAAAGTTTCAATATTTGTGAGGAAAAACAAAGCTTAGGCTGTTAAATAAAAGCTATTCTATTTTGG
40 TAAAGTTGAAATTCAGAAATGTATATATAATTTCTTAAATGTAAGAGTGCAATTAATCAATAACAAAAATGGTTTTTCAATGCA
TCATTATTTATTGGCATAACAAATTTGTTGGTCCAAAAAGTAAAGCTATATTCAGAAAGTTTGCACCTCAAAATTCAGTAATAAT
TAATGAAATGCTGTACAAAAATAAGTGAAGCAGTGT

3

45 MQMRRVWRTCRSGPRKRREARYKKSTEDIFPAQLLKLQRHERVWQEPVDRHRSWGGSGAGGVAGREWTDQGGVALGGHYMAE
GEGYFAMSEDELACSPYIPLGGDFGGDFGGDFGGDFGGGSGGGHCLDYCESPTAHCNVLNWEQVQRLDGLSETIPIHGR
GNFPTLEQLPSLIVKVVRRRLAEKRIGVRDVRNLGSAASHVLHQDSGLGYKDLDLIFCADLRGEGEFQTVKDVLDCLDLFLPE
GVNKEKITPLTLKEAYVQKMKVVCNDSRWLSLSNNSGKNVELKFVDSLRRQFEFSVDSFQIKLDSLLLFYECSENPMETET
HPTIIGESVYGDFQEAFLCNKIIATRNPBEIRGGGLLKVCNLLVRGFRPASDEIKTLQRYMCSRFFIDFSDIGEQQRLSEY

LQNHFVGLDRKYEYLMTLHGTVNESTVCLMGHERRQTLNLITMLAIRVLADQNVI PN VANVT CYYQ PAPYVADANFSNYIIAQ
VQPVFTCCQQTYSTWLPNCN

4

5 ACAACTGCTAAAGCTCCAGAGACACGAGCGTGTGTGGCAGCAAGAGCCGCCAGTTCGGGACCACCGCAGCTGGGGTGGCAGCGG
CGCAGGAGGGGTGCGGGGCGGGAGTGGTGAGCGCAGGCGGAGGGGTCTGGGAAAGACGAAGTCGCTATTTGCTGTCTGAGCG
CGCTCGCAGCTCCTGGAAGTGTTCGCCCTCTCGGTTCGCTCTCGCTCGCTCGCTCCTAGAGGGGCGGCGCCTCCAGGAC
TGACCAGGGCCAAGTGGCGCTCGGCGGGCACTACATGGCGGAGGGTGAAGGGTACTTCGCCATGTCTGAGGACGAGCTGGCCTG
CAGCCCCACATCCCCCTAGGCGGGCACTTCGGCGGCGGCGACTTCGGCGGCGGCGACTTCGGCGGCGGCGACTTCGGCGGCGG
CGACTTCGGCGGTGGCGGCGAGCTTCGGTGGGCATTCGCTGGGACTATTGGCGAAAGCCCTACGGCGCACTGCAATGTGTGAACTG
10 GGAGCAAGTGACGCGGTGGACGGCATCTGAGTGAGACCATTCGGAATTCACGGGCGCGCAACTTCGCCACGCTCGAGCTGCA
GCCGAGCCTGATCGTGAAGGTGGTGGCGGCGGCGCTGGCGGAGAAGCGCATTGGCGTCCGCGAGCTGGCGCTCAACGGCTCGGC
AGCCAGCCATGTCTGCACAGGACAGCGGCTGGGCTACAAGGACCTGGACCTCATCTTCTGCGCCGACCTGGCGGGGAAGG
GGAGTTTCAGACTGTGAAGGACGTCGTGTGGACTGCTGTGGACTTCTTACCCGAGGGGGTGAACAAAGAGAAGATCACACC
ACTCAGCTCAAGGAAGCTTATGTGCAGAAAATGGTTAAAGTGTGCAATGACTCTGACCGATGGAGTCTTATATCCCTGTCAAA
15 CAACAGTGGCAAAAATGTGGAACGAAATTTGTGGATTCCCTCCGAGGCGAGTTGAATTCAGTGTAGATTCTTTTCAAATCAA
ATTAGACTCTCTCTGTCTTTTATGAATGTTTACAGAGAACCAATGACTGAGACATTTACCCCCACAATAATCGGGGAGAGCGT
CTATGGCGATTTCAGGAAGCCTTTGATCACCTTTGTAACAAGATCATTTGCCACCAGGAACCCAGAGGAAATCCGAGGGGAGG
CCTGCTTAAGTACTGCAACCTCTTGGTGAGGGGCTTTAGGCCCGCTCTGATGAAATCAAGGCCCTTCAAAGGTACATGTGTTC
CAGGTTTTCATCGACTTCTCAGACATTTGAGAGCAGCAGAGAAAATGGAGTCTTATTTGCAGAACCACTTTGTGGGATTGGA
20 AGACCGCAAGTATGAGTATCTCATGACCCCTTCATGGAGTGGTAAATGAGAGCACAGTGTGCCTGATGGGACATGAAAGAAGACA
GACTTTAAACCTTATCACCATGCTGGCTATCCGGGTGTTAGCTGACCAAAATGTCAATTCCTAATGTGGCTAATGTCACTTGGCTA
TTACCAGCCAGCCCCCTATGTAGCAGATGCCAATTTAGCAATTACTACATTGCACAGGTTACGCCAGTATTACGTGCCAGCA
ACAGACCTACTCCACTTGGCTACCTGCAATTAAGAATCATTTAAAAATGTCTGTGGGAAGCCATTTTCAGACAAGACAGGAG
AGAAAAAAAAAAAAAAAAAAAA

25 5

MKKFSRMPKSEGGSGGGAAGGAGAGAGAGCGSGSSVGVRFVAVGRHQVTLEESLAEGGFSTVFLVRTHGGIRCALKRMVYN
NMPDLNVCKREITIMKELSGHKNIYGLDCAVNSISDNVWEVLILMEYCRAGQVNVQMNKKLQTFTEPEVLQIFCDTCEAVAR
LHQCKTPIIHRDLKVENILLNDGGNYVLCDFGSATNKFNPQKDGNNVVEEIKKYTTLSYRAPEMINLYGKPIITTKADIWAL
GCLLYKLCFFTLPFGESQVAICDGNFTIPDNSRYSRNIHCLIRFMLEPDPEHRPDIQVSYFAFKFAKKDCPVSNINKCKQLL
30 RHGALLTEILLFLQLFLNRMTASEAAARKSQIKARITDTIGPTETSIAPRQRPKANSATTATPSVLTIQSSATPVKVLAPGEFG
NHRPKGALRPGNGPEILLQGPPQPPQQRVLLQQLQGDWRLQQLHLQHRHPHQQQQQQQQQHNNHHHHLLQDAYMQOYQ
HATQQQQLQOQFLMHSVYQPPSASQYPTMMPQYQQAFFQQQMLAQHQPSQQQASPEYLTSPQEFSPALVSYSLSLPAQVGTI
MDSSYSANRSVADKEA IANFTNQKNI SNPPMSGWNPFGE DNFSKLTEBELL DREFDLRSSKGHLKAYPASQ

6

35 AGCCGGGCGAGCTGCAGCGGAGCCGCGGAGCGGGCGGCGGGGCCAGGCTGTGCGCTTGGGAGCGCGGAATGTGAGGCTTGGCG
GGCCGACGACGCTCGGACGGGCGGAGGGGCGGCGACCCCTCGCGGACGCGCGGCTGCGCGCGGGCGGCGGCGGACTTGCCTTGC
CGCTCCCTGCGCCTCCAGCTCGCCGCGGGGACCATGAAGAAGTCTCTCGGATGCCCAAGTCGGAGGGCGGCGAGCGGCGGCGG
AGCGGCGGGTGGCGGGGCTGGCGGGGCGGGGCGGGGCGGCTGCGGCTCGCGGCGGCTCGTCCGTGGGGGTCCGGGTGTTCGC
GGTCGGCGGCCACAGGTCACCTGGAAGAGTCGCTGGCCGAAGGTGGATTCTCCACAGTTCCTCGTGGCTACTCACGGTGG
40 AATCCGATGTGCATTGAAGCGAATGTATGTCAATAACATGCCAGACCTCAATGTTGTGAAAAGGAAATTACAATTATGAAAGA
GCTATCTGGTCACAAAAATATTGTGGGCTATTTGGACTGTGCTGTAAATTCAAATTAGTGATAATGTATGGGAAGTCCTTATCTT
AATGGAATATTGTGAGCTGGACAGGTAGTGAATCAAAATGAATAAGAAGCTACAGACGGGTTTTACAGAACCAGAAGTGTACAC
GATATTCTGTGATACCTGTGAAGCTGTTGCAAGGTGTCATCAGTGAAGACTCCAATAATTACCGGGATCTGAAGGTAGAAAA
TATTTTGTGTAATGATGGTGGGAATATGTACTTTGTGACTTTGGCAGTGCCACTAATAAATTTCTTAATCCTCAAAAAGATGG
45 AGTTAATGTAGTAGAAGAAGAAATAAAAAGTATACAACCTCTGTATACAGAGCCCTGAAATGATCAACCTTTATGGAGGAA
ACCCATCACCACCAAGGCTGATATCTGGGCACTGGGATGTCTACTCTATAAACTTTGTTTCTTCACTCTTCCTTTTGGTGAGAG
TCAGGTGTCTATCTGTGATGGCAACTTACCATCCCAGACAATTTCTGTTACTCCCGTAACATACATTGCTTAATAAGGTTTCAT
GCTTGAACCAGATCCGGAACATAGACCTGATATATTTCAAGTGTATATTTGCATTTAAATTTGCCAAAAAGGATGTCCAGT
CTCCAACATCAATAAGTGTGTGAAACAATTACTGAGACACGGAGCACTGTTAAGTGAATTTCTTCTATTCTTCACTCTTCCT

ATTTCTCTCTGAAGAGGGAGGGCAGGACCAAGGGGAAAGTCTGTGTCACAGGCCGAAACCTCATGAGTTAAAGAGTCATCAACT
CAGCTCCAACCTGCAATCGTGACCTGATGAGAGGAGGTACTCCCTGCCTTAAAGAGCATCTACATGCGACAGATTGACCAGAT
GGTTGGTTTGAATTTGATCTGCTCTGTGAGGCCCTCATCTTGAACAGACCTTAAACGACCAAACCCAGACCATGCCAAAGTCCA
GTCTGAAATGAAAGGGGACAGAAGTCTTGCTGTCCATGGTGGTGTAGGAATTTCTGTGCAACACCTCACCACGCTCTGGTTAATC
5 CTTGCACACTTCAGTGTTTCTCTCTCTGTTCAATAAAATGCCCTGTTAAGGATATAATTTGAAGTGAGAAGATACATGGAAATTG
CCCTCTTATGACATGTTGATGTTATAAGCACAATAGATGGGGCATCTTTGGATTGATGTTACAGCTTTATATCTCAGAACCTA
AGTCTCTTCACTTTGCCTGGCACCTGCTATACTGGAGTATTGCTATGCTCTTTAAAAATTTTTTTTTTATATATTTATTTTTTT
GAGACAGGGTCTTGATATTTTTTTGGGACAGGGTTACCTGGGCTCAAGTGATCCTTCTGCCTCAGCCTCCCGAGTAGCTGGGAT
TACAGGTGAGCACCACGTACCTGGCTAGCTACTTCTTTGTTAGAGGATTGAGAAATGAAATTTCTGCAAAAGGGCCCATGGTTT
10 ATTTGGTATCCCTATTTAATTGCATTGAAAATGTCATCCTTTCTGTTGTTAGATAAATGGGGTCTTCCCTGATATCCAACCGT
GATTTTGGATCACATGGGAGAAAAAGTCATCCAGTTTTTCATGTTTGCCTCAAGTAATCTTTACAGTGTTACAAATTATTTGCT
TAAGAAGAATGGTCTTAAACAGAATTCTTAACAGATAGTCTCTTAGGTTATTATGTTATGGTCTAAGAGGTTAACTGACATCTT
TTGGATGGTATTTTGCATTTTGAATATGAACCTTACCTGAGGAACCTCCCATAGTTCCAGAATCAGGTGCCTTTTTAGGGAGAGAAC
AATACCTAAGATTGCTGAGCTTCCATCTTTCTCATATTTCTTAAGCAAGGATCTCACTTATGACCATATTTGGGTTAGAGTT
15 CTGTTTTGTTTTCTGTTTTCTGTGCTAGTGCCAATTAGCTAAATCAGGAGAAAGAAATGATCACATGACTTTTAGCATCCTTG
AGCCATTTCTCTGTGTAATACAGGCTTTAGATTAGTCCCTGATATTTGGTTTTGGTTTTGGGGCAGTGGATGTCGCAGCTATGCT
ATGGTTTTCAGGAGGCTGTTTTAGCCACATGGTGAGACCGTTGGTGAAGGGGGATGGAATTCGTTGGCCAGCTCTTGCCCTTTCA
TCTGTAAAGTAAGCATGTAGAAGAGGAAGTTGTGCTTAAATGCCCTTTGTTTTTTTTTGTATTATTTTCTAGCCAGAACATC
TCTCTTTGAACFCACACTGATACACACCTGCTACTCTTACACAGTGCAGCAGGGCTGACTCTTTAGTCTGGCTTCCATGAAGCGT
20 CATGGGTGGAACGCATTTCTAGTAAAAAAGGTAGGAAATCCCTAAAACTTCCAGCCTCATATAGCAGGTTCTCACCTGTCACT
GTTTTCCACCTCTAAGGATTTTCATGTACATCTTTTCAAAGCTAGAAATAAGCACTGTCTAAGTTTATGTTGCATTTTTAGTCA
AAAGGGAGAAATCTTATCTCTTCTGAAAAATTTAAGTGTTATGGTTTTATATAGTTTCACTTCTTTGAGATTTTTGAAAAGAGT
ATTTTCAGTAATAAACGTGCCATCTCTATCTCTTAAACATTTATTAACAATTTGTTTTAAAAATAGAAAAATAAAATGCTTCT
ATTTTACCTTTTTTCATTTTCAAGCATTATTTCTGTTTTATTACAGTGCTCCATCTACTGAATAGAAAATTTGAGAATAATAT
25 ATATATATATTTTAAATGTTTTCACTGACTCATTGAAAATGTTAATTACACACATGCATGCATGCACACAGCAGCATACTTG
TACCTTTTGTCTCTGGGCAACAGGTGGGACTGTTAGTGACCCATTTGGGAAAAATAGAGCATCTCAGAGAAGGAGGTGAGTTCTT
CCTGCCTGTGATTTCTCTTGGGCTCCCTCCTCTCCGCTCTGGCTTCTGTGGCGGCAGTGTTGGGTAGGCACTCCAGTGTTT
TCTTAATGAGGCACTTTGCTGTCACTCGAGCAAGCCTGGGTGTTCTTCTCTCTCATGCTCCTGGAATAGGGAATAGGGATCT
CATGCTTGCAAACTACACAATGTGTCAGGTGCTTCCCAGGGGCCACAGGCTGTGAGGAAACGTGTTTTATGTTAAGTCACAAAC
30 CCACTTGACTTCTGGGTACTGGAATTAATACAGTGGGTGAGACTGAGGTTGAGTGAGTTAGTACATATTAATCCTGGTTGTTG
AGCTTCCAGACTACCCCGTCCAAAGTTTGATGCTATGTAGTCACTGTTGTTGGGGCTGGATGCCAGAAGGTTCTTTGAGCCAG
TTTCAAAGGTTACTTGTTTTTTTTTTTTTTTTTTTTTTTAAAGTCAGAATGTTAACAGCTGTGATATATCTGACAGGCTTTTGCAGTT
TCTTCTGTCTGTGTTCTGAAATCTGGGTAGAGAATGGCTGAGGAGGAGATTACCGAGAAGTTGTTTGTCTGCTCAGTGTCTTG
CCAGGATTTGCCCTCAAACTGAGTGGACTTCATCCTTTGGCGCGCTCTGAGCCTGGCCATCTTCTTATTTCCCGAGCTGTAGCT
35 AGTGTCTAGTGTGAGCTTTGCTCAATGTGGTGGAAACATTTTGCAGAATGTTGTAGAAAGCTGCCTTATAGTTGGCTTGACAA
AGCATAATTTCTCTCATACAAACTTTCAAATCATTTACAGTAGCTTAGCTACTTTAGTTGATGTGACCGAGGAATCCCTTCTAGA
ATCATAGGTGGCAAGGGAGGGTTTGCTAGCTCTCCATTTGCACCTGGCCATTTGTGAAAAACCAGCTTCTGTATTCAAATCTTTCC
TTTCAATTTTTTAAATTTTTTTTTTTTGGCAGCGCTGTGCTGGAACCTTACTCATTTGTAACCTGAATCCTCAGGCTTTTCTGTGTTT
AGATCATGGACTGTGCACGTGACACTTAAATAATTTTCTATGTATTTAAAGAAAAATGCACCAGGATGGTGTCTGTGCACGTGA
40 CTATTAGAGGAGCGTCTGTAGAAGTACCTGGTTTGGTCAGTGCAGTTGTGCAATCTGAGGCGCTTGTTCCTCTCCCTTTCC
CCTTCTCCCCACCAAAGGAAAAATATCCCTCTTAATGATTTTCGTAGTTTCACTTACTGAATGATTACCACCTGTAATCTCTCTTT
GGATTGTGTAGACTCAACATGAGACATTCCTTTCTGCTTTCTGGAGGGCACCAGGGGCTTTCTCTTTGATAAAATTTTTTTTGT
CTGTTGACAAAAACAAAAATCTTTTTCAAATGTAGTGTGTTGAAAAGGTAGGCTGAGTGATTTACCTTAGCCACAGGGTGGC
TGAGCAGGAACCTTTAGAAGAAAACTCTGAGCTTTCTGTGCCATTTCCAGCATCCAGCTCCTATTCTAGTGCCCTCTTCCCTGCAG
45 GGCAGGGACCCCTTGGGAAATCGAGGAGGTGGGACGGGCTGGGCCCTGTGTCCAGGTTTTCACAGGGCTCAGGGTTATGCTCCC
GCTTGAATCTGGACGTGAATCTGGTAAAAATATCAAGTACCTGTGGAACCTCCCTGATTCTATACCTCTTCTCTCTTCTGCAA
GGCAGAGGAATAATATTTTTTAAAGGTTATTTTGTTTTAGTTTTTAAATAGCAAAACACAAGCTGCATTTTTATTTATTTTGCATA
AGAAAGGTAAATCTTTTTTACAAAAAAAAGTATAGAGTTGGAACTCTGGGAAACCTTACGGAATACACAAATGCTTCTCTGTA
ATGTGCAATATGCTTTGCAACTGTAGATGATATTTATGTTTAACTGTAAATAAGAAATGATTTTAAATTAAAGGGATCTTT
50 TTGTAAAGGACCAATGTTCTTTTATAAATGTAATAAGGAATATCTTGCTCTTTAAATTTATTAGGATTTTATGAGTAATTT
TTTATTAAGAAATTTCTTTTTTTTTT

9

MACEIMPLQSSQEDERPLSPFYLSAHVPQVSNVSATGELLERTIRSAVEQHFLFDVNNSSGGQSSSEDSSEGTLSASSATSARQRRR
QSKEQDEVHRGRDKGLINKENTPSGFNHLDDCILNTQVEKVKHNTFGCAGERSKPKRQKSSTKLSELHDNQDGLVNMESLNST
RSHERTGPDDFEWMSDERKGNEKDGGHTQHFESPTMKIQEHPSLSDTKQQRNQDAGDQEESEFVSEVPQSDLTALCDEKNWEEPI
5 PAFSSWQRENSDSDEAHLSPQAGRLIRQLLEDESDPMLSPRFYAYGQSRQYLDDEVPSPSPNSHSPMRRSSSLGSDYDDEQED
LTPAQLTRRIQSLKKKIRKFEDRFEEKKYRPSHSDKAANPEVLKWTNDLAKFRRLKESKLKISEEDLTPMRQRNLTLPKSF
GSQLEKEDEKKQELVDKAIKPSVEATLESIQRKLQEKRAESSRPEDIKDMTKDQIANEKVALQKALLYESIHRPVTKNERQV
MKPLYDRYRLVKQILSRANTIPIIGSPSSKRRSPLLQPIIEGETASPFKEIKEEEGSEDDSNVKPDFMVTCLKTDFSAFCFLDQ
FEDDADGFISPMDDKIPSKCSQDTGLSNLHAASIPELLEHLQEMREEKKRIRKKLRDFEDNFRQNGRNVQKEDRTPMAEEYSE
10 YKHIKAKRLRLEVLISKRDTSKSM

10

AAATGTAGAGAAGCAGCCGATAAAATAGCATTGCCTGAAGAAGTTTGGAGGCTGAGAGCAGCAGTAGACTGGCCAACTGCAGAG
CAAGTTGTTCCTCCAGCCGTGCGGTGCAGCCTCATGCCCCCAACCCAGCTTAGCCACTGTAAGAAGACGTTCACTGTACAGACG
ACCAAACCTGCCGTGGAAGAGACAGTTGTGAGATTCCCTTGCAAATTTACATACGAGAATGGCTTGTGAAATCATGCCTCTGCA
15 AAGTTCACAGGAAGATGAAAGACCTCTGTCACCTTTCTATTGTAGTGCTCATGTACCCCAAGTCAGCAATGTGTCTGCAACCGG
AGAACTCTTAGAAAGAACCATCCGATCAGCTGTAGAACAACATCTTTTGTAGTGAATAACTCTGGAGGTCAAAGTTCAGAGGA
CTCAGAATCTGGAACACTATCAGCATCTTCTGCCACATCTGCCAGACAGCGCCGCCAGTCCAAGGAGCAGGATGAAGTTCG
ACATGGGAGAGACAAGGGACTTATCAACAAGAAAATACTCCTTCTGGGTTCACACCTTGATGATGTATTTGAATACTCA
GGAAGTCGAAAAGGTACACAAAATACTTTTGGTGTGTCTGGAGAAAGGAGCAAGCCTAAACGTCAGAAATCCAGTACTAAACT
20 TTCTGAGCTTCATGACAATCAGGACGGTCTTGTGAATATGGAAGTCTCAATTCACACGATCTCATGAGAGAACTGGACCTGA
TGATTTTGAATGATGTCTGATGAAAGGAAAGGAAATGAAAAGATGGTGGACACACTCAGCATTTTGGAGCCCCACAATGAA
GATCCAGAGCATCCCAGCCTATCTGACACCAACAGCAGAGAAATCAAGATGCCGGTGACCAGGAGGAGAGCTTTGTCTCCGA
AGTGCCCCAGTCGGACCTGACTGCATTTGTGTGATGAAAAGAAGTGGGAAGAGCCTATCCCTGCTTCTCCTCTGGCAGCGGGA
GAACAGTGACTCTGATGAAGCCCACCTCTCGCCGAGGCTGGGCGCCTGATCCGTCAGCTGCTGGACGAAGACAGCGACCCCAT
25 GCTCTCTCCTCGGTTCTACGCTTATGGGCAGAGCAGGCAATACCTGGATGACACAGAAGTGCTCCTTCCCCACCAAACCTCCCA
TTCTTTCATGAGCGCGGAAGCTCCTCTCTGGGTCTTATGATGATGAGCAAGAGGACCTGACACCTGCCAGCTCACACGAAG
GATTCAGAGCCTTAAAAAGAAGATCCGGAAGTTTGAAGATAGATTGGAAGAAGAGAAGATACAGACCTTCCCACAGTGACAA
AGCAGCCAAATCCGAGGTTCTGAAATGGACAAATGACCTTGCCAAATTCGGGAGACAACTTAAAGAATCAAACTAAAGATATC
TGAAGAGGACCTAACTCCAGGATGCGGCAGCGAAGCAACACACTCCCCAAGAGTTTGGTTCCCACTTGAGAAAGAAGATGA
30 GAAGAAGCAAGAGCTGGTGGATAAAGCAATAAAGCCCAAGTGTGGAAGCCACATTGGAATCTATTCAGAGGAAGCTCCAGGAGAA
GCGAGCGGAAAGCAGCCGCCCTGAGGACATTAAGGATATGACCAAGACAGATGCTAATGAGAAAGTGGCTCTGCAGAAAGC
TCTGTTATATTATGAAAGCATTATGAGCGGCCGGAACAAAGAACGAGGAGGATGAAGCCACTATACGACAGGTACCG
GCTGGTCAAAACAGATCTCTCCGAGCTAACACCATACCCATCATTTGGTTCCCTCCAGCAAGCGGAGAAGCCCTTTGCTGCA
GCCAATTATCGAGGGCGAACTGCTTCTCTTCAAGGAGATAAAGGAAGAAGAGGAGGGGTGAGAAGCAGATAGCAATGTGAA
35 GCCAGACTTCATGGTCACTCTGAAAACCGATTTCAGTGACAGATGCTTTCTGGACCAATTCGAAGATGACGCTGATGGATTAT
TTCCCCAATGGATGATAAAATACCATCAAAATGCAGCCAGGACACAGGGCTTTCAAATCTCCATGCTGCCTCAATACCTGAACT
CCTGGAACACCTCCAGGAAATGAGAGAAGAAAAGAAAGGATTGGAAGAACTTCGGGATTTTGAAGACAACCTTTTCAGACA
GAATGGAAGAAATGTCAGAAGGAAGACCGCACTCCTATGGCTGAAGAATACAGTGAATATAAGCACATAAAGCGGAACTGAG
GCTCCTGGAGGTGCTCATCAGCAAGAGAGACACTGATTCCAAGTCCATGTGAGGGGCATGGCCAAGCAGGGGGCTGGCAGCT
40 GCGGTGAGAGTTTACTGTCCCCAGAGAAAGTGACGCTCTGGAAGGCAGCCTTGGGGCTGGCCCTGCAAAGCATGCAGCCCTTCT
GCCTCTAGACCATTGGCATCGGCTCCTGTTTCCATTGCCTGCCTTAGAAACTGGCTGGAAGAAGACAATGTGACCTGACTTAG
GCATTTTGAATTTGAAAGTCAAGACTGCAGTATGTGCACATGCGCACGCGCATGCACGCACACACACACAGTAGTGAGGCT
TTCTTAACACTAGCAGAGATTAATCACTACATTAGACAACACTCATCTACAGAGAATATACACTGTTCTTCCCTGGATAACTGA
GAAACAAGAGACCAATTCTCTGTCTAACTGTGATAAAACAAGCTCAGGACTTTATTTCTATAGAGCAAACTTGTGTGGAGGGCC
45 ATGCTCTCCTTGGACCCAGTTAACTGCAACGTGCATTGGAGCCCTATTGTCTGCCGCTGCCATTTCTAGTGACCTTTCCACAGA
GCTGCGCCTTCCCTACGTGTGTGAAAGGTTTCCCTTCAGCCCTCAGGTAGATGGAAGCTGCATCTGCCACAGATGCGAGTGC
AGTCATCATCTTCAGGATGTTCTTCAGGACTTCTCAGCTGACAAGGAATTTTGGTCCCTGCTAGGACCGGGTCACTCTGCAG
AGGACAGAGAGATGGTAAGCAGCTGTATGAATGCTGATTTTAAACACAGGTCATGGGAGAAGAGCTTGAGATTTCTTCTGAA
CACTGACTGCATTTACCAGTCTGATTTTATCGTCAAACACCAAGCCAGGCTAGCATGCTCATGGCAATCTGTTTGGGGCTGTTT
50 TGTGTGGCACTAGCCAAACATAAAGGGGCTTAAGTCAGCCTGCATACAGAGGATCGGGGAGAGAAGGGGCTGTGTTCTCAGC
CTCCTGAGTACTTACCAGACTTTAATTTTAAAAAAAATCTGCACTAAAATCCCCAACTGACAGGTAAATGTAGCCCTCAG

AGCTCAGCCCAAGGCAGAATCTAAATCACACTATTTTCGAGATCATGTATAAAAAGAAAAAAGAAAGTCATGCTGTGTGGCCA
ATTATAATTTTTCACAAAGACTTTGTGCACAAACTGTCTATATTAGACATTTTGGAGGGACCAGGAAATGTAAGACACCAAATC
CTCCATCTCTTCAGTGTGCCTGATGTACCTCATGATTGTCTGTACTTTTAACTCCTGCGCCAAGGACAGTGGGTCTGTG
TCCACCTTTGTGCTTTGCGAGGCCGAGCCAGGCATCTGCTCGCTGCCACGGCTGACCAGAGAAGGTGCTTCAGGAGCTCTGC
5 CTTAGACGACGTGTACAGTATGAACACACAGCAGAGGCACCCCTCGTATGTTTTGAAAGTTGCCTTCTGAAAGGGCACAGTTTT
AAGGAAAAGAAAAGAAATGTAAACTATACTGACCCGTTTTCAGTTTAAAGGGTCTGTGAGAACTGGCTGGTCCAATGGGATT
TACAGCAACATTTTCCATTGCTGAAGTGAGGTAGCAGCTCTCTTCTGTGCTGAGTGAATGTTAAGGATGGGGAAAAAGAAATGCCCTT
TAAGTTTGTCTTTAATCGTATGGAAGCTTGAGCTATGTGTTGGAAGTGCCCTGGTTAATCCATACACAAAGACGGTACATAATC
CTACAGGTTTAAATGTACATAAAATATAGTTTGAATCTTTTGTCTTACTGTTTACATTGCAGATTGCTATAATTTCAAGGAG
10 TGAGATTATAAATAAAATGATGCACCTTTAGGATGTTTCTATTTTGAATCTGAACATGAATCATTACATGACCAAAAAATTG
TGTTTTTTTAAAAATACATGCTAGTCTGTCTCTTAAATAGCTCTCTTAAATAAGCTATGATATTAATCAGATCATTACCAGTTA
GCTTTTAAAGCACATTTGTTTAAAGACTATGTTTTTGGAAAAATACGCTACAGAAATTTTAAAGCTACAAATAAATGAGATG
CTACTAATGTTTGAATCTGTTGTTTCTGCCAAAGGTAATTAACATAAGATTTATTCAGGAATCCCCATTTGAATTTGTAT
GATTCATAAAAGAAAAACCAAGTAAGTTATATAAAAT

15 11
MLKGVQRKRADKYWEYTFKVNWSDLSVTTVTKTHQELQEFLKLKPKELSSETFDKTLRLALNQSLKREERRHPDLEPILRQLF
SSSSQAFQLSQKVSFFQSISSDSLHSINLQSSLKTSKILEHLKEDSSEASSQEEDVLQHAIIHKHKTGKSPIVNNIGTSCSP
LDGLTMQYSEQNGIVDWRKQSCCTTIQHEHCVTSDQHSAEKRLSSINKKKGKPQTEKEKIKKTDNRLNSRINGIRLSTPQHA
HGGTVKDVNLDIGSGHDTGSETSESYSPPSPRHGDRSFESSEKDRDTSNEDSGNPSTTRFTGYGSVNQTVTVKPPVQI
20 ASLGNENGNLLEDPLNSPKYQHISFMPTLHCVMHNGAQKSEVVVPAPKPADGKTIGMLVPSVAISAIRESANSTPVGILGPTA
CTGESEKHELLELASPLIPSTFLPHSSTPALHLTVQRLKLPPPQGSSECTVNIPOQPPGSLSIASPNATAFIPIHNPSPFPSP
VATTDPIPKSASQVVLNQMVPIEGNTCTVPQPTNVKVLPAAGLSAAQPPASYPPLPGSPLAAGVLPSONSSVLSTAATSPQP
ASAGISQAQATVPPAVPHTPGPAPSPSPALHTSTAQSDSTSYISAVGNTNANGTVVPPQMGSGPCGSCGRRSCGCTGNLQL
NSYYYPNMPGPMYRVPSTFLPSICNGSYLNQAHQSNGNQLPFFLPQTPYANGLVHDPVMGSGANYGMQMGAGFRFYPVYPA
25 PNVVANTSGSGPKNGNVSCYNGVSGHYAQDCKQSSMEANQQGTYRLRYAPPLPPSNDTLDSDAD

12
ATCGGAAAAGCACCTTGAGTTACTGGCTTCCCTTTTACCTATTCCATCAACCTTCCTTCCACACAGTAGTACTCCCGCTTTGCA
TCTTACAGTTTACAGGGCTAAAGTTGCCACCACACAGGGATCTTCTGAGAGCTGCACAGTTAACATCCCACAACAACCCCGG
AAGCCTGAGCATCGCATCAACAAACACTGCCTTTTATCTTATCCATAACCCAGGTAGTTTCCAGGCTCTCCTTGCATACCAC
30 GGACCCCATCACAAAATCTGCATCCCAAGTGGTAGGACTCAATCAAAATGGTGCCCTCAAATTGAGGGAACACAGGGACAGTCCC
TCAGCCTACCAATGTGAAGGTAGTTCTTCCAGCAGCTGGCCTCTCAGCTGCTCAGCCACCAGCTTCCCTACCCCTTACCAGGCTC
TCCCTTGTGCGCGCGTGTACCAGCCAGAACTCCAGTGTGCTCAGCACAGCAGCAACTTCTCCCCAGCCAGCGAGCGCAGG
TATCAGCCAGGCCAGGCAACTGTTCCTCCTGCAGTTCTTACCCACACCCAGGCCCTGCCCCGAGCCCAAGCCCTGCCTTGAC
ACACAGTACCGCGCAGAGTGACAGCACCTCTTACATCAGTGTGTTGGGGAACACGAACGCTAATGGGACAGTAGTGCCACCGCA
35 GCAGATGGGCTCAGGTCTTGTGGTTCTTGTGGGCGAAGGTGCAGCTGTGGGACCAATGGAACCTTCAGCTAAATAGTTACTA
TTATCTTAATCCAATGCCGACCAATGTACCGAGTCCCTTCATTCTTTACTCTGCCATCCATTGCAATGGCAGCTACCTCAA
CCAAGCACATCAGAGCAATGGAACCAACTTCTTTTTTCTGCCTCAGACTCCATATGCAATGGACTGGTACATGACCCAGT
CATGGGGAGCCCAAGCCAACTATGCGATGCAGCAGATGGCAGGATTGGGAGATTCTATCTGTATATCCAGCACCTAACGTAGT
TGCCAACACCAAGTGGTTCGGGGCCCAAGAAGAAATGGGAATGTCTCATGTTACAATGTGGTGTAAGCGGACACTACGCACAGGA
40 CTGTAAGCAGTCGTCATGGAGGCAATCAACAAGGCACCTACAGACTGAGATACGCACCTCCCTCCCTTCTAATGATAC
GTTGGATTCTGCAGACTGAAACGAGTAAAGCTTGCTACTTAATACACTCAAGTGTGGGAGTCATGGGGTGTGGAGGGAGGA
AAGGAAAGGTATTTGTCTTCTTGTCTATACATTTTCTAGATTTCTATGCAAGTTGGGATTTTTCATTCTCTTGTACCAATGTC
CAAAACAAGAAAGAAATGCAATGCTTTTGAGCCTCTGGTCTCCTGGTTCAACAACAGGCTTATATGTATGATACATGTAATTTAA
ACCTTCAGACAACTTAAATGTTGGTGCGTGCTTTTTTTTTTTTACTGAATACTTGCTGTGTGAATGTTTACTGAATCT
45 TTAAGCTGTGATTTGACCTTTTTTTTACAACACTGGTGACAGTCATATGGTTTTGAAAAAAAAGAAATTTTGCTTCTTCC
CAGCTTTTCTCATTTCACCTTAAACGACACTTCTCCCCAGCCAGCCTCAGCTCTGTCTCCGGCCCGCAGCAGGAGCAGCCAGC
AGTGCAATTCACCCACTTTGTAACTGCTCTGCATATAAACAAGGGCAGAATGTTTACCCTGATCTTATGGGAGGAATCGA
ACTCCAAAATAGTGTGATATATGTAATAAACAGCGTCACGTAATACATATATGCAGTGTCTTGTGTTCCAAATAGAAAATGAA
AATAAGTGAAGAGAGAGGAAGAAGTCAACCATATGAAACTGAAAAAATATGACGTACGAAATGGACAAAAGCTTTTTCTGAA
50 ACCAACTTTTACTTCCATCATCTTTTTTAGCCTGTGCTTCAGAGAGACACAAAGTGAACACACTGGTGTGAATGTGCTCT

CTGTGTGCTTGTGTTTGTAAATGAAAGTCTACAGCCAATTTTACTTGTCTACCACCGTGTGTGCTCAAAGAGACACTACTTGAG
TGAAGATTCTTCTTTCCCTGTACCAGCTGTACAGTGTACGTTGTGTTTAAATGTGTATGGTTTATTGCAATCTGAACAGA
GCTATGGGTTTCTACCATAAGTCAGGTTGTTTGTCCCTAACCTGTCTCTCGTAGCAAAGTCACTTTTATAACAGTTTACCCT
ATGCTTGATPATAATGTGAAAGGCGGAATTCAGTGTGTTAAGATGGTATTATCATGTCCGTGTCATGTCTACTAAGTTTAAAT
5 GCTGCTGTTTAAAAAAGTTTAAAAAGCCAATCTATGTACTAAATTCCTTCCAGGTAATTTTIGATTTCCTA
AAGTGCAGTGAAGTTATCTGGAAGATTGGGTGTATTTTGGTGACTGCTGCATTTCATCAGCAATGAACAGTTTCCACTGTATA
GTCCTAGGGGTGAGGGGTGGGGTTTCATTTCCATTCCTCAGCACAGAGCAGAAATGATAGATTATTATTGTTTGGAGTAAC
GTTGGTATGCAGCAGAGGAACATAAACAATTTGGTCTTGGTTCAGAAGCCTAACAGATTGCTAGACAAGAGAAAACTTGAAGA
AAAAAGAAGCTTAATTTTCATGCTTCATAAGTAGCATTTATATTTATAGCACCATGTACATTTTGAACTTTCTTTCAGGGTG
10 GGAGTTATGGGAAGGGGTGGGTGTGAAGGGTAGATGAAAGCTTAAATTTAGAAAGAAAGTTCAAGTAAAGGAAATTTATTTG
ATTAAATATATTTTATTGATCTGGGTATTTTGGACCACATTTATAAATTAATTTGTTAAGTGCAGTTGAGTTGTTCAAGTGA
GAGTTTGTATAAGCCACTTATGGCCCGCTTGTGAATCACTTGCAGTTGTACTTTATGGAGCTTATTTTATGATTTAAATAC
TGTAAGTATACATAGGAGGTATGTTACCTTCTCCTTATTTGTATGTTTACCATACTTTGATATTTGAAATGTTATGTACTGGA
AAGGCCACTTATATTTCTAGAACAGATTGGATTTTATGCAACCTTTTTCCTTGAATTAACAGCAATAAAAAATGAAAAACAA
15 AAAAAAAAAAAAAAAAAA

13

RIFGASPPRNLAIQKASRTAAAMGSEDHGAQNPSCKIMTFRPTMEEFKDFNKYVAYIESQGAHRAGLAKIIPPKEWKPRQTYD
DIDDDVPIPAIQQVVTGQSLFTQYNIQKKAMTVGEYRRLANSEKYCTPRHQDFDLERKYWKNLTFVSPFIYADISGLYDDD
VAQWNIGSLRTILDMVERECGTIEGVNTPYLYFGMWKTTFAWHTEDMDLYSINYLHFGEPKSWYAIPEHGKRLERLAIGFFP
20 GSSQGCDAFLRHKMTLISPIILKKYGIPIFSRITQEGEFMITFPYGYHAGFNHGFNCAESTNFATLRWIDYKGVATQCTCRKDM
VKISMDVFRILQPERYELWKQKDLTVLDHTRPTALTSPELSSWSASRASLAKALLRRSHRKSQPKPKPEDPKFPCEGTAG
AALLEEAGGSVKEEAGPEVDPEEEEEEPQLPHGREAEAEEDGRGKLRPTKAKSERKKKSFGLLPPLPPPAHFPSSEALWL
PSPLEPPVLGPGPAAMEESPLPAPLNVVPEVPSELEAKPRPIIPMLYVVP RPKAAPNQHVSQQAFEHFAQKGPTWKEPV
SPMELTGPEDGAASSGAGRMETKARAGEGQAPSTFSKLMEIKKSRRHPLGRPPTRSPLSVVQKQASSDEEASPFSGEEDVSDP
25 DALRPLLSLQWKNRAASFQAEKFNAAAARTEPYCAICTLFYPYQALQTEKEAPIASLGEGCPATLPSKSRQKTRPLIPEMCF
TSGGENTEPLPANSYIGDDGTSPLIACGKCLQVHASCYGRPELVNEGWTCSRCAAHAWTAECCLCNLRGGALQMTTDRRWIH
VICAIAPPEARFLNVIERHPVDISAIPEQRWKLKCVYCRKRMKVSGACIQCSYEHCTSFHVTCAHAAGVLMPEDDWPVVS
TCLKHKSGGHAVQLLRVSLGQVITKNRNGLYYRCRVIGAASQTCYEVNFDGSGSYSDNLYPESITSRDCVQLGPPSEGELEVEL
RWTDGNLYKAKFISSVTSHIYQVEFEDGSQTLVKRGDIFLLEELPKRVR SRLSLSTGAPQEPAPSGEEAKAARPRVGTPLAT
30 EDSGRSQDYVAFVESLLQVQGRPGAPP

14

GTCGCCAGCAACCGAGCGGGGCACGCCGAGCGGGGCTGGGGGTGCGAGCCGAGGGCGGGGAGAGCGCGCCGCTGCTCCCGG
ACCGGGCCGCGCACGCCGCTCAGGAACCATCACTGTTGCTGGAGGCACCTGACAAATCTAGCGAATTTTGGAGCATCTCCA
CCCAGGAACCTCGGCATCCAGAAGTGTGCTTCCGCGACAGCTGCAGCCATGGGGTCTGAGGACCAGCGGCCAGAACCCAGC
35 TGTAATCATGACGTTTCGCCCAACCATGGAAGATTTAAAGACTTCAACAAATACGTGGCTACATAGAGTCGAGGAGCC
CACCGGCGGGCCTGGCCAAGATCATCCCCGAGGAGTGAAGCCGCGGCAGACGTATGATGACATCGACGAGCTGGTGATC
CCGGCGCCCATCCAGCAGGTGGTGACGGGCCAGTCGGGCTCTTACAGCAGTACAATATCCAGAAGAAGGCCATGACAGTGGGC
GAGTACCGCCGCTGGCCAACAGCGAGAAGTACTGTACCCCGCGGCACCGAGCTTTGATGACCTTGAACGCAATACTGGAAG
AACCTCACCTTTGTCTCCCGATCTACGGGGCTGACATCAGCGGCTCTTTGTATGATGACGAGCTGGCCAGTGAACATCGGG
40 AGCCTCCGACCATCTGGACATGGTGGAGCGGAGTGGCGCACCATCATCGAGGGCGTGAACACGCCCTACCTGTACTTCGGC
ATGTGGAAGACCACCTTCGCTGGCACACCGAGGACATGGAACCTGTACAGCATCAACTACCTGCATTTGGGGAGCCTAAGTCC
TGGTACGCCATCCACAGAGCACGGCAAGCGCTGGAGCGGCTGGCCATCGGCTTCTTCCCGGGAGCTCGCAGGGCTGCGAC
GCCCTCCTCGCGCATAGATGACCTCATCTCGCCCATCATCTGAAGAAGTACGGGATCCCCTTACAGCGGATCAGCAGGAG
CGCGGGGAATTCATGATCATATTTCCCTACGGCTACACCGCGGCTTCAATCAGGGTTCAACTGCGCAGAATCTACCAACTTC
45 GCCACCTGCGGTGATGACTACGCCAAAGTGGCCACTCAGTGACGTCGCGGAAGGACATGGTCAAGATCTCCATGGACGTG
TTCGTGCGCATCTGCAGCCCGAGCGCTACGAGCTGTGGAAGCAGGGCAAGGACCTCACGGTGTGACACACCGGGCCACG
GCGCTACACAGCCCCGAGCTGAGCTCCTGGAGTGATCCCGGGCTCGCTGAAGGCCAAGCTCCTCCGAGGTCTACCGGAAA
CGGAGCCAGCCAAAGAAGCCGAAGCCGAAGACCCCAAGTTCCCTGGGGAGGTACGGCTGGGGCAGCGCTCTAGAGGAGGCT
GGGGGAGCGTGAAGGAGGAGGCTGGGCCGAGGTTGACCCGAGGAGGAGGAGGAGGAGCGGAGCCACTGCCACAGGCCGG
50 GAGGCCGAGGGCGCAGAAGAGGAGCGGAGGGCAAGCTGCGGCCAACCAAGGCCAAGAGCGAGCGGAAGAAGAGCTTCGGC

CTGCTGCCCCACAGCTGCCGCCCCGCTGCTCACTTCCCCCTCAGAGGAGGCGCTGTGGCTGCCATCCCCACTGGAGCCCCCG
GTGCTGGGCCCCAGGCCCTGCAGCCATGGAGGAGAGCCCCCTGCCGGCACCCCTTAATGTCGTGCCCCCTGAGGTGCCAGTGAG
GAGCTAGAGGCCAAGCCTCGGCCCATCATCCCCATGCTGTACGTGGTGCCGCGCGGGCAAGGCAGCCTTCAACCAGGAGCAC
GTGTCTGCCAGCAGGCCCTTTGAGCACTTTGCCCAGAAGGGTCCGACCTGGAAGGAACCACTTTCCCCCATGGAGCTGACGGGG
5 CCAGAGGACGGTGCAGCCAGCAGTGCGGCGAGGTCGCATGGAGACCAAGCCCGGGCCGGAGAGGGGCAGGCACCGTCCACATTT
TCCAAATTTGAAGATGGAGATCAAGAAGAGCCGGCGCCATCCCCCTGGGCGGGCCGCCACCCGGTCCCCACTGTGCGTGGTGAAG
CAGGAGGCCCTCAAGTGACGAGGAGGCATCCCCCTTTCTCCGGGAGGAAGATGTGAGTGACCCGGACGCCCTTGAGGCCGTGTG
TCTCTGCACTGGAAGAACAGGGCGGCCAGCTTCCAGGCCGAGAGGAAGTTCAACGCAGCGGCTGCGCGACGGAGCCCTACTGC
GCCATCTGCACGCTCTTCTACCCCTACTGCCAGGCCCTACAGACTGAGAAGGAGGCACCCATAGCCTCCCTCGGAAAGGGCTGC
10 CCGGCCACATTACCCCTCCAAAAGCCGTGAGAAGACCCGACCGCTCATCCCTGAGATGTGCTTACCTCTGGCGGTGAGAACAG
GAGCCGCTGCGCTGCCAACTCCTACATCGGGCAGCAGGGACAGCCCCCTGATCGCTGCGGCAAGTGCTGCCTGCAGGTCCAT
GCCAGTTGTGATGGCATCCGTCCCGAGCTGGTCAATGAAGCTGGACGTGTTCCTGGTGCGCGGCCACGCCGTGGACTGCGGAG
TGCTGCCCTGTGCAACCTGCGAGGAGGTGCGCTGCAGATGACCAGGATAGGAGGTGGATCCACGTGATCTGTGCCATCGCAGTC
CCCGAGGCGCGCTTCTGAACTGATTTAGCGGCCACCCCTGTGGACATCAGCGCCATCCCCGAGCAGCGGTGGAAGCTGAAATGC
15 GTGTACTGCGGAAGCGGATGAAGAAGGTGTGAGGTGCTGTATCCAGTGCTCCTACGAGCACTGCTCCACGCTCTCCACGTG
ACCTGCGCCACGCCGAGGCGTGTCTATGGAGCCGACGACTGGCCCTATGTGGTCTCCATACCTGCCTCAAGCACAAAGTCG
GGGGGTACGCTGTCCAACTCCTGAGGGCCGTGTCCCTAGGCCAGGTGGTTCATCACCAGAACCAGAACGGGGTGTACTACCGC
TGTGCGCTCATCGGTGCGCCCTCGCAGACCTGTACGAAGTGAAGTTCGACGATGGCTCCTACAGCGACAACCTGTACCCTGAG
AGCATCACGAGTAGGGACTGTGTCCAGCTGGGACCCCCCTTCGAGGGGGAGCTGGTGGAGCTCCGGTGGACTGACGGCAACCTC
20 TACAAGGCCAAGTTTCATCTCTCCGTACCCAGCCACATCTACCAGGTGGAGTTTGAGGACGGGTCCAGCTGACGGTGAAGCGT
GGGGACATCTTACCCCTGGAGGAGGAGTGCCTAAGAGGGTCCGCTCTCGGCTGTCACTGAGCACGGGGGCACCGCAGGAGCCC
GCCTTCTCGGGGGAGGAGGCCAAGGCCGCCAAGCGCCCGGTGTGGGCACCCCGCTTGGCCACGGAGGACTCCGGGCGGAGCCAG
GACTACGTGGCTTTCGTGGAGAGCCTCCTGCAGGTGCAGGGCCGGCCCGGAGCCCCCTTCTAGGACAGCTGGCCGCTCAGGCGA
CCCTCAGCCCGGGCGGGAGGCCATGGCATGCCCGGGCCCTTCGCTTGTGTGAATTCCTGTCTCGTGTCCCCGACCCCGAGA
25 GGCCACCTCCAAGCCGCGGGTGGCCCCCTAGGGGACAGGAGCCAGCGGGACGCCGCACGCGGCCAGACTCAGGGAGCAGGGC
CAGGCGGGCTCGGGGGCCGGCCAGGGGAGCACCCCACTCAACTACTCAGAATTTTAAACCATGTAAGCTCTCTTCTTCGAAA
AGGTGCTACTGCAATGCCCTACTGAGCAACCTTTGAGATTGTCACTTCTGTACATAAACCACCTTTGTGAGGCTCTTTCTATAA
ATACATATTTGTTTAAAAAAGCAAGAAAAAAGGAAAAAAGGAAAAATATCCCCAAGTTGTTTCTAGATTTGTGGCTTTA
AGAAAAACAAAACAAAACACATTGTTTCTCAGAACAGGATTTCTCTGAGAGGTGAGGCATCTCGCTGTTTCTTTGTT
30 GTTGTTTTAAAAATATTGATTTGGCTACAGACCAGGCGAGGAAAGAGACCCGGTAATTGGAGGGTGAACCTCGGGGGGGGGC
AGGACGCCCCGGTTTCGGCACAGCCCGGTCACTACGGCTCGCTCTCGCTCACCCCGGCTCTGGCTTTGATGGTGTGCTG
CCAGTGCCCTGTGCCCACTCTGTGCTGTGCGGAGGAGGCCAGGCTCTCTGGTGGCCGGCCCTGTGCACCTGGCCAGGGGAAGC
CCGGGGGTCTGGGGCTCCCTCCGTCTGCGCCACCTTTGCAGAATAAACTCTCTCTGGGGTTGTCTATCTTTGTTTCTCTC
ACCTGAGAGAAACGCAGGTGTTCAGAGGCTTCTTGCAGACAAAGCACCCCTGCACCTCCTATGGCTCAGGATGAGGGAGGCC
35 CCCAGGCCCTTCTGGTTGGTAGTGAGTGTGGACAGCTTCCAGCTCTTCCGGGTACAACCTGAGCAGGTGCGGGGACACAGGGC
CGAGGCAGGCCTTCGGGGCCCTTTTCGCTGTCTCCGGGACAGGACAGGCCTGGTGTCTCTGCTCCACCCACCCACGCTGCTG
TCACCTGAGGGGAATCTGCTTCTTAGGAGTGGGTGAGCTGATAGAGAAAAACGGCTTCAGCCAGGCTGGGAAGCGCTTC
TCCAGGTGCTCTCCCTCACCAGCTCTGCACCCCTCTGGGGAGCCTTCCACCTTAGCTGTCTCTGCCCCAGGGAGGGATGG
AGGAGATAATTTGCTTATATTAAAAACAAAAATGGCTGAGGCAGGAGTTTGGGACAGCTTGGCTATATAGCAAGACCCCAT
40 CACTACAAATTTTACAAATTAGCTAGGTGTGTTGGTGGCGACCTGTGGTCCAGCTACTCGGAGGCTGTGGTGGGAGGATT
GCTTGAGTCCAGGAGGTGAGGCTGCAGTCACTGACCACTGCACTCCAGCTGGGCAACAGAGCGAGACCTGTCT
CCAAAAAAGGCAATGTTTATATTATAAAGAGTGTCTAACAGTCCCCGGGCTAGAGAGGACTAAGGAAAACAGAGA
GAGTGTACGCAGGAGCAAGCCTTTCATTTCCTTGGTGGGGAGGGGGCGGTTGCCCTGGAGAGGGCCGGGGTCCGGGAGGTT
GGGGGGTGTGAGCCAAACGTGGAGGTGTCCCTCTGCACGCAGCCCTCGCCCGGCTGGCGCTGACACTGTATTTCTATGTTGT
45 TTGAAAATGCTATTTATATTGTAAAGAAGCGGGCGGGTGGCCCTGTGCTTGTCCCTTGGGGGTACACCCATCCCCTGGT
GGCTCTGGGCGGCTGCGCAGATGGGCCACAGAAGGGCAGGCCGAGCTGCACACTCTCCCAACGAAGGTATCTGTGTCTT
ACTCTGTGCAAGACGCGGCAAAACCCAGTGCCCTGGTTTTCCTCCCGAGATGAAGGATACGCTGTATTTTGTGCTAATG
TCCCTGCCCTTAGGTTTCAATGAATTAAGGTTTCAAGCGCTGCG

15

50 MTTTATVDYDNIETQYSDVNNRWDDWDNENSSARLFERSRIKALADEREAVQKKTFTKWNHSLARVSCRITDLYTDLRD
GRMLIKLEVLSEBRLPKPTKGRMRIHCLNVKALQFLKEQVRHLENMGSHDIVDGNHRLTLGLIWTIILRFQIQDISVETED

NKEKKSADALLWCQMKTAGYPNVNIHNFTTSWRDGMFANALIHKHPDLIDFDKLKSSNAHYNLQNAFNLAEQHLGLTKLLD
 PEDISVDHPDEKSIITYVVTYYHYFSKMKALAVEGKRIGKVLDAIETEKMIKEYESLASDLEWIEQTIIILNNRKFANSLVG
 VQQQLQAFNTYRTVEKPPKFTEKGNLEVLFTIQSKMRANQKVMPREGKLI SDINKAWERLEKAEHERELALRNLIRQEKL
 EQLARRFDRKAAMRETWLSQNLVSQDNFGFDLPAVEAATKKHEAIEDIAAYEERVQAVVAVARELEAENYHDIKRITARKD
 5 NVIRLWEYLLELLRARRQRLEMNGLQKIFQEMLYIMDWMDEMKVVLVLSQDYGKHLGVEDLLQKHTLVEADIGIQAERVGVN
 ASAQKFATDGEYKPCDPQVIRDRVAHMEFCYQELCQLAAERRARLEESRRLWKFWEAEEEGWIREKEKILSSDDYGKDLTS
 VMRLLSKHRAFEDEMSGRSGHFEQAKEGEDMIAEEHFGSEKIRERIIYIREQWANLEQLSAIRKKRLEEASLLHQFQADADDI
 DAWMLDILKIVSSSDVGHDEYSTQSLVKKHKDVAEEIANRPTLDTLHEQASALPQEHAEPSDVRGRLSGIEERYKEVAELTRL
 RKQALQDTLALYKMFSEADACELWIDEKEQWLNMMQIPEKLEDLEVIQHRFESLEPEMNNQASRVAVVNQIARQLMHSGHPSEK
 10 EIKAQQDKLNTRWSQFRELVDKRDALLSALSQNYHLECNETKSWIREKTKVISTQDLGNDLAGVMALQRKLTGMERDLVAI
 EAKLSDLQKEAEKLESEHPDQAQAILSRALAEISDVWEEMKTTLNREASLGEASKLQOFLRDLDDFQSWLSRTQTATASEDMPN
 TLTEAEKLLTQHENIKNEIDNYEEDYQKMRDMGEMVTQGTDAQYMFRLRQLQALDTGWNELHKMWENRQNLSSQSHAYQQFLR
 DTKQAEAFNNQEVYLAHTEMPTTLEGAAAIKKQEDFMTTMDANEKINAVVETGRRLVSDGNINSRIQEKVDSIDDRHRKN
 RETASELLMRLKDNRLQKFLQDCQELSLWINEKMLTAQMSYDEARNLHSHKWLKHQAFMAELASNKWEWDKIEKEGMQLISEK
 15 PETEAVVKEKLTGLHKMWVLESTTQTKAQRFLDANKAELFTQSCADLDKWLHGLEBSQIQSDDYGKHLTSVNILLKKQOMLENQ
 MEVRKKEIEELQSAQALSQEGKSTDEVDSKRLTVQTKFMELLEPLNERKHNLASKEIHQFNRDVEDEILWVGEMPLATSTD
 HGHNLTQVQLLIKNNQTLQKEIQGHQPRIDDI FERSQNIIVTSSSLSAEAIQRRLADLKQLWGLLIEETEKHRRLLEEAAHRAQQ
 YYFDAEAEAWMSEQELYMSEEKAKDEQSAVSMKKKHQILEQAVEDYAETVHQLSKTSRALVADSHPESEIRISMRQSKVDKLY
 AGLKDLAEERRGKLDERHRLFQLNREVDDLEQWIAEREVWAGSHELQGDYEHVMTLQERFREFARDTGNIGQERVDTVNHLADE
 20 LINSCHSDAATIAEWKDLNEAWADLLELIDTRTQILAASYELHKFYHDAKEIFGRIQDKHKKLPEELGRDQNTVETLQRMHTT
 FEHDIQALGTQVRQLQEDAAARLQAYAGDKADDIQKRENEVLEAWKSLLDACESRRVRLVDTGDKFRFFSMVRDLMLWMDVIR
 QIEAQEKPRDVSSVELLMNNHQGKAEIDARNDSTTTCIELGKSLARKHYASEEIKEKLLQLTEKRKEMIDKWEDRWEWLRLI
 LEVHQFSRDASVAEAWLLGQEPYLSREIGQSVDEVEKLIKREAFKSAATWDERFSALERLTLELLEVRQEQEEERKRPP
 PSPEPSTKVSEEAESQQQWDTSKGEQVSQNLPAEQGSPRMAETVDTSEMVNGATEQRTSSKESPIPSPTSDRKAKTALPAQS
 25 AATLPARTQETPSAQMEGFLNRKHEWEAHNKKASSRSWHNVYCVINNQEMGFYKDAKTAASGIPYHSEVPVSLKEAVCEVALDY
 KKKKHVFKLRLNDGNEYLFQAKDDEEMNTWIAISSAISSDKHEVASTQSTPASSRAQTLPTS SVVTITSESSPGKREKDKED
 KEKRFSLFGKKK

16

GTGAGCTGAAGCAGGGCAGGGCATCAACTACCCAGGAAGTGCAAGGGGTTTGGGGATTTTCCTTTCTTAGCCAAGGGAAGGCA
 30 TGACAGACTGTACTGGAAAAACAGGACACTCTTGCCCAAATACCTGCACTTTTGCACAGCTCTTAGCAACTGGCAGACCAGGAG
 ATTCCTCTCTGTGCTTCAATCATTGGGTCCACACCCATAGGGCCCTTGCTTACTGCCAGTGCAGCAGCTGAGATTAACACCCC
 ATCCCCGGGAGAACTCTAAGAAGGAGCTGATGTGGAGGAGCAGCTGAGACAGTTCAAGATGACGACACAGTAGCCACAGACTA
 TGACAACATTGAGATCCAGCAGCAGTACAGTGATGTCAACAAACCGCTGGGATGTGACGACTGGGACAATGAGAACAGCTCTGC
 GCGGCTTTTGTAGCGGTCCCGCATCAAGGCTCTGGCAGATGAGCGTGAAGCCGTGCAGAAGAAGACCTTACCAAGTGGGTCAA
 35 TTCCACCTTGCCCGTGTGCTGCTGCGGATCACAGACCTGTACACTGACCTTCGAGATGGACGGATGCTCATCAAGCTGTGGA
 GGTCTCTCTGGAGAGAGGCTGCCATAACCCACCAAGGGACGAATGCGCATCCACTGCTTAGAGAAATGTGGACAAGGCCCTTCA
 GTTCTGAAGGAGCAGAGAGTCCATCTTGAGAACATGGGGTCCCATGACATCGTGGATGGAAACCACCGGTGACCTTGCCCT
 CATCTGGACCATCATCTTGGCTTCCAGATCCAGGATATCAGTGTGGAACATGAAGACAACAAGAGAAGAAATCTGCCAAGGA
 TGCATTGCTGTTGTGGTGCCAGATGAAGACAGCTGGGTACCCCAATGTCAACATTCAAAATTCACCACTAGCTGGAGGGACGG
 40 CATGGCCTTCAATGCACTGATACACAAACACCGGCTGACCTGATAGATTTTGACAAACTAAAGAAATCTAACGCACACTACAA
 CCTGCAGAAATGCAATTAATCTGGCAGAACAGCACCTCGGCCTCACTAAACTGTTGGACCCGGAAGACATCAGCGTGGACCATCC
 TGATGAGAAAGTCCATAATCACTTATGTGGTGACTTATTACCACTACTTCTCTAAGATGAAGGCCTTAGCTGTTGAAGGAAAACG
 AATTGGAAGGTGCTTGACAAATGCTATTGAAAACAGAAAAATGATTGAAAAGTATGAATCACTTGCCCTTGACCTTCTGGAATG
 GATTGAACAAACCATCATCTTCTGAACAAATCGCAAATTTGCAATTCACCTGGTGGGGTTCAACAGCAGCTTCAGGCATTCAA
 45 CACTTACCGCACTGTGGAGAAACACCCAAATTTACTGAGAAGGGGAACCTTGAAGTGCTGCTTTCACCAATTCAGAGCAAGAT
 GAGGGCCAAACACAGAAAGTCTACATGCCCCGGGAGGGGAAGCTCATCTCGACATCAACAAGGCCTGGGAAAGACTGGA
 AGCGGAACACGAAAGAGAACTGGCTTTGCGGAATGAGCTCATAAGACAGGAGAACTGGAACAGCTCGCCCGCAGATTTGATCG
 CAAGGCAGCTGAGGGGAGACTTGGCTGAGCGAAACACAGCGTCTGGTGTCTCAGGACAACCTTTGGGTTTGACCTTCTCGTGT
 TGAGGCCCGGCACAAAAAGACAGGCGCATTTGAGACAGCACTTGGCCGATACGAGGAGCGTGTGACGCTGTGGTAGCCGTGGC
 50 CAGGGAGCTCGAGGCGGAGAAATTACCACGACATCAAGCGCATCACAGCGAGGAAGGACAATGTCATCCGGCTCTGGGAATACCT
 ACTGGAACCTGCTCAGGGCCCGGAGACAGCGCTCGAGATGAACCTGGGGCTGCAGAAGATATTCAGGAAATGCTCTACATTAT

GGACTGGATGGATGAAATGAAGGTGCTAGTATTGTCTCAAGACTATGGCAAACACTTACTTGGTGTGGAAGACCTGTTACAGAA
GCACACCTTGGTTGAAGCAGACATTGGCATCCAGGCAGAGCGGGTGAGAGGTGTCAATGCCCTCCGCCAGAACTTCGCAACAGA
CGGGGAAGGTTACAAGCCCTGTGACCCCCAGGTGATCCGAGACCGCGTGGCCACATGGAGTTCTGTTATCAAGAGCTTTGCCA
GCTGGCGGCTGAGCGCAGGGCCCGTCTGGAAGAGTCCCGCCGCTCTGGAAGTTCTTCTGGGAGATGGCAGAAGAGGAAGGCTG
5 GATACGGGAGAAGGAGAAGATCCTGTCTCGGACGATTACGGGAAAGACCTGACCAGCGTCATGCGCTGCTCAGCAAGCACCG
GGCGTTCCAGGACGAGATGAGCGGCCGAGTGGCCACTTTGAGCAGGCCATCAAGGAAGGCGAAGACATGATCGCGGAGGAGCA
CTTCGGGTCGGAGAAGATCCGTGAGAGGATCATTTACATCCGGGAGCAGTGGGCCAACCTAGAGCAGCTCTCGGCCATTCGGAA
GAAGCGCTGGAGGAGGCTCCCTGCTGCACCACTTCCAGGCAGATGCTGATGACATTGATGCCTGGATGCTGGACATCCTCAA
GATTGTCTCCAGCAGCGACGTGGGCCACGATGAGTATTCACACAGTCTCTGGTCAAGAAACACAAGGACGTGGCGGAAGAGAT
10 CGCCAATTACAGGCCACCCCTTGACACGCTGCACGAACAAGCCAGCGCCCTCCCCAGGAGCATGCCGAGTCTCCAGACGTGAG
GGGACGGCTGTCCGGCATCGAGGAGCGGTATAAGGAGGTGGCAGAGCTGACGCGGCTGCGGAAGCAGGCACTCCAGGACACTCT
GGCCCTGTACAAGATGTTTCAGCGAGGCTGATGCCTGTGAGCTCTGGATCGACGAGAAGGAGCAGTGGCTCAACAACATGCAGAT
CCCAGAGAAGCTGGAGGATCTGGAGGTTCATCCAGCACAGATTTGAGAGCCTAGAACCAGAAATGAACAACAGGCTTCCCGGT
TGCACTGGTGAACCAGATTGCACGCCAGCTGATGCACAGCGGCCACCAAGTGAGAAGGAAATCAAAGCCCAGCAGGACAACT
15 CAACACAAGGTGGAGCCAGTTTCAGAGAACTGGTTGACAGGAAGAAGGATGCCCTCCTGTCTGCGCTGAGCATCCAGAACTACCA
CTTCGAGTGCAATGAAACCAAATCCTGGATTGCGGAAAGACCAAGGTCATCGAGTCCACCCAGGACCTGGGCAATGACCTGGC
TGGCGTTCATGGCCCTGCAGCGCAAGCTGACCGGATGAGCGGGACTTGGTGCCATTTGAGGCAAGCTGAGCTGACCTGACGAGAA
GGAGGCGGAGAAGCTGGAGTCCGAGCACCCGACAGGCCAGGCCAGGCCATCCTGTCTCGGCTGGCGGAGATCAGCGACGTGTGGGA
GGAGATGAAGACCACCTTGAAAAACCGAGAGGCTCCCTGGGAGAGGCCAGCAAGCTGCAGCAGTTCTACGGGACTTTGGACGA
20 CTTCAGTCTTGCTCTTAGGACCCAGACAGCGATCGCCTCGGAGGACATGCCAAACACCTGACCGAGGCTGAGAAGCTGCT
CACGCAGCAGGAGAATCAAGAATGAGATCGACAACACGAGGAGGACTACCAGAAGATGAGGACATGGGCGAGATGGTCAC
CCAGGGGACAGCCGATGCCAGTACATGTTCTGCGGACGCGCTGCAGGCCCTGGACACTGGATGGAACGAGCTCCACAAGAT
GTGGGAGAACAGACAAAATCTCCTATCCAGTCACATGCCTACCAGCAGTTCTTCAGAGACACGAAGCAAGCCGAAGCCTTTCT
TAACAACAGGAGTATGTTCTGGCTCACACTGAAATGCCTACCACCTTGAAGGAGCTGAAGCAGCAATTAAGCAAGAGGA
25 CTTTCATGACCACCATGGAGCGCAATGAGGAGAAGATCAATGCTGTGGTGGAGACTGGCCGGAGGCTGGTGAGCGATGGGAACAT
CAACTCAGATCGCATCCAGGAGAAGGTGGAATCTATTGATGACAGACATAGGAAGAATCGTGAGACAGCCAGTGAACCTTTGAT
GAGGTTGAAGGACAACAGGGATCTACAGAAATTCCTGCAAGATTGTCAAGAGCTGTCTCTTGATCAATGAGAAGATGCTCAC
AGCCAGGACATGTCTTACGATGAAGCCAGAAATCTGCACAGTAAATGGTTGAAGCATCAAGCATTTATGGCAGAACTTGCATC
CAACAAAGAATGGCTTGACAAAATCGAGAAGGAAGGAATGCAGCTCATTTTCAGAAAAGCCTGAGACGGAAGCTGTGGTGAAGGA
30 GAAACTACTGGTTTACATAAAATGTGGGAAGTCTTGAATCCACTACCAGACAAAGGCCAGCGGCTCTTTGATGCAAAACAA
GGCCGAACCTTTTACCCAGAGCTGTGCAGATCTAGACAAATGGCTGCACGGCTGGAGAGTCAAGTTCAGTCTGATGACTATGG
CAACACCTGACCAAGTGTCAATATCCTGCTGAAAAGCAACAGATGCTGGAGAATCAGATGGAAGTGGCGAAGAAGGAGATCGA
AGAGCTCCAAGCAAGCCAGGCCCTGAGTCAGGAAGGGAAGACACCGAGGTCAGACAGCAAGCGCCCTACCGCTGCAGAC
CAAGTTTCATGGAGTTGCTGGAGCCCTTGAACGAGAGGAAGCATTAACCTGCTGGCTTCAAAGAGATCCATCAGTTCAACAGGGA
35 TGTGGAGGACGAGATCTTGTGGGTGGAGAGAGATGCCTTTGGCAACTTCCACGGATCATGGCCACAACCTCCAGACTGTGCA
GCTGTTAATAAAGAAAAATCAGACCCCTCCAGAAAGAAATCCAGGGGCACAGCCTCGCATTTGACGACATCTTTGAGAGGAGCCA
AAACATCGTCACTGACAGCAGCAGCTCAGCGCTGAGGCCATCAGACAGAGGCTTGCCGACCTGAAGCAGCTGTGGGTCTCCT
CATTTGAGGAGACAGAGAAACGCCACAGGCGGCTGGAGGAGGCGCACAGGCCACAGCAGTACTACTTTGACGCTGCTGAGGCCGA
AGCCTGGATGAGCGAGCAGGAGCTGTACATGATGTCAGAGGAGAAGGCCAAGGATGAGCAGAGTGTGTCTCCATGTTGAAGAA
40 GCACCAGATCTTAGAACAAAGCTGTGGAGGACTATGCAGAGACCGTGCATCAGCTCTCCAAGACCAGCCGGGCCCTGGTGGCCGA
CAGCATCTTGAAAGTGAGCGCATTAGCATGCGGAGTCCAAGTGGAATAAAGTGTACGCTGGTCTGAAAGACCTTGCTGAAGA
GAGAAGAGGCAAGCTGGATGAGAGACACAGGTTATTCCAGCTCAACCGGAGGTGGACGACCTGGAGCAGTGGATCGCTGAGAG
GGAGGTGGTCGAGGTTCCATGAACTGGGACAGGACTATGAGCATGTACGATGTTACAAGAAGCATTCGGGAGTTTGGCCG
AGACACCGGAACATTGGGAGGAGCGCGTGGACACGGTCAATCAGCTGGCAGATGAGCTCATCAACTTGGACATTGAGATGC
45 CGCCACCATCGCTGAATGGAAGGATGGCCCTCAATGAAGCCTGGGCCGACCTCCTGGAGCTCATTTGACACAAGAACACAGATTCT
TGCCGCTTCTATGAACCTGCACAAGTTTACCACGATGCCAAGGAGATCTTTGGGCGTATACAGGACAAACACAAGAACTCCC
TGAGGAGCTTGGGAGAGATCAGAACACAGTGGAGACCTTACAGAGAATGCACACTACATTTGAGCATGACATCCAGGCTCTGGG
CACACAGGTGAGGACGCTGCAGGAGGATGAGCCCGCTCCAGGCGGCTATGCGGGTGACAAGGCCGACGATATCCAGAAGCG
CGAGAACGAGGTCTTGAAGCCTGGAAGTCCCTCTGGACCGCTGTGAGAGCCGAGGTGCGGCTGGTGACACAGGGGACAA
50 GTTCCGCTTCTTACGATGGTGGCGGACCTCATGCTCTGGATGGAGGATGTCATCCGGCAGATCAGGCCCCAGGAGAAGCCAA
GGATGTATCATCTGTTGAACCTCTTAATGAATAATCATCAAGGCATCAAGCTGAAATGATGCACGTAATGACAGTTTCAAC
CTGCATTGAACCTGGGAAATCCCTGTTGGCGAGAAAACACTATGCATCTGAGGAGATCAAGGAAAAATTTACTGCAGTTGACGGA

AAAGAGGAAAGAAATGATCGACAAGTGGGAAGACCGATGGGAATGGTTAAGACTGATTCGGAGGTCCATCAGTTCCTCAAGAGA
CGCCAGTGTGGCCGAGGCTTGGCTGCTTGGACAGGAGCCGTACCTATCCAGCCGAGAGATAGGCCAGAGCGTGGACGAGGTGGA
GAAGCTCATCAAGCGCCACGAGGCATTTGAAAAGTCTGCAGCAACCTGGGATGAGAGGTTCTCTGCCCTGGAAAGGCTGACTAC
ATTGGAGTTACTGGAAGTGCAGCAGACAGCAAGAGGAAGAGGAGAGGAAGAGGCGGCCCTTCTCCGAGCCGAGCAGCAAGGT
5 TTCAGAGGAAGCCGAGTCCCAGCAGCAGTGGGATACTTCAAAAGGAGAACAAGTTTCCCAAAACGGTTTGCCAGCTGAACAGGG
ATCTCCACGGATGGCAGAAACGGTGGACACAAGCGAAATGGTCAACGGCGCTACAGAACAAAGGACGAGCTTAAAGAGTCCAG
CCCCATCCCCCTCCCGACCTCTGATCGTAAAGCCAAGACTGCCCTCCAGCCAGAGTGCCGCCACCTTACCAGCCAGAACCCA
GGAGACACCTTCGGCCAGATGGAAGGCTTCCCTCAATCGGAAACACGAGTGGGAGGCCACAAAGAAAGCCTCAAGCAGGTC
CTGGCACAATGTTTATGTGTCTATAAATAACCAAGAAATGGGTTTCTACAAAGATGCAAGACTGCTGCTTCTGGAATTCCTTA
10 CCACAGCGAGGTCCCTGTGAGTTTGAAAGAAGCTGTCTGCGAAGTGGCCCTTGATTACAAAAGAAGAAACAGTATTCAAGCT
AAGACTAAATGATGGCAATGAGTACCTCTTCAAGCCAAAGCAGATGAGGAAATGAACACATGGATCCAGGCTATCTCTTCCGC
CATCTCTCTGATAAACACGAGGTGTCTGCCAGCACCCAGAGCAGCCAGCATCCAGCCGCGCAGACCCCTCCCCACACGCGT
CGTCACCATCACCAGCGAGTCCAGTCCCGCAAGCGGGAAGGACAAAGAGAAAGCAAAGAGAAGCGGTTTACGCTTTTGG
CAAAAAGAAATGAATCCTTCTCTCACCTCTGCCCCCTCTCTTACCTTTTTCAGTGAATTCACAGCATGCAAGCTCAGAACCAA
15 CACATTACTCTCTGTGCTAATGTCTCAATGTGGTGTGATTTATTTTATTTTAAATTTATAGAGCATTTCCGGGGGGGTGGG
G
17
MQPDLKQILPFPLEPAPTLGLFSNYSTMDPVQKAVLSHTFGGPLLTKRPVISCNICQIRFNSQSQAHAHYKGNRHARRVKG
IEAAKTRGREPVREPGDPAPPGSTPTNGDGVAPRPVSMENGLGPAGSPEKQPGSPSPPSIPETGQGVTKGEGGTAPASLPG
20 GSKEEEEAKRLLYCALCKVAVNSLSQLEAHNKGTKHKTTLEARSGLPIKAYPRLPPTPGEPEAPQDRTFHCEICNVKNS
EVQLKQHISSRRHRDGVAGKPNLLSRHKSRGAGELAGTLTFSKELPKSLAGLLPSPLAVAAVMAAAGSPLSLRPAPAPL
LQGPPIITHPLLHPAPGPIRTAHGPILFSPY
18
AGAGCGACGCGGAACCCCGGGCGCTGGGTCCCCAGCATGATCCTCGGCAGCCTGAGCCGGGCGAGGCCCTGCCTCTGCTACG
25 GCAGCCCCGATCATGCAGCCCCACTGGACCTCAAGCAGATCCTGCCCTTCCCACTCGAGCCAGCCCCACCTTGGCCTCTT
CAGCAACTACAGCACCATGGACCTGTGCAGAAGGCTGTGCTCTCCCACTTTTGGGGACCTTGTCAAGACCAAGCGACC
CGTCATTTCTGTAAATATCTGTCAAATCCGCTTCAATTTCTCAGAGCCAGGCTGAGGCGCACTACAAAGTAAATCGCCACGCCG
ACGAGTCAAAGGCATTGAGGCTGCAAGACCAGAGGCGAGGAGCCTGGCGTCCGAGAACCTGGAGACCCAGCTCCCCAGGCAG
CACCCCAACAAATGGGGATGGTGTAGCACCCCGTCCAGGTACTAAGCACAAGACAATTCTGGAGCCCGAAGTGGGCTCGGGCC
30 CATCAAAGCTTACCCTCGGCTGGGGCCTCCACCCCGGGGAACAGAGGCTCCTGCCAGGACCGAACTTTCCACTGTGAGAT
CTGCAATGTCAAAGGTCAACTCGGAGGTCCAATGAAACAGCACATCTCCAGCCGGCGGACCGAGACGGCGTGGCCGGGAAGCC
CAACCCACTACTGAGCCGTCACAAGAAGTCTAGGGGCGCGGGGAGCTGGCGGGCACGCTGACTTTCTCCAAGGAGCTGCCCAA
GTCCCTGGCGGGCGGCCTGCTCCCCAGCCCCCTGGCGGTGGCTGCAGTATGGCAGCGGCAGCAGGCTCGCCGCTGTCCCTGCG
CCCGGCTCCAGCCGACCTCTTCTCCAGGACCGCGGATCACTACCCCTCTGCTTACCCCGCCCCCGGACCATCCGAATGC
35 GCACGGACCCATCCTCTTCTCCCGTACTGACCTCAACCCTGAACCCCTCCCATTCAACTCCCCACCTCCAGCCGGGACCCAGG
CGTCCGACTCCAGCCCGCCCCCTCTCCCGGACTCCCTAAATGATCTCTCTTCCCCCCCCCACCAGATACGGGGTTC
CAGGAAAGGGGAGGGGTAGCGGGGAGGGGGTTCAGAAGGGGGGAACACCCAGATCTCAGGGAACCCCGCCCCCTGCCTT
TCCCTCTCCCCTAGAAAAGGGGGGCGCTCTACCCCGAGCCCCCTTGGAGACACCCCCCCCTCCCAAAAGCCATGTCCATCCA
GCCCTTCCCCCAAACCTAGCACAAAACGGGTTTCAAGCCATGGTTCGGGTCGGGGGGACAGAAATGGATTTTCTTGGCA
40 ATAAGCGGACTCTGGGACTCCGGCTCCCTACCCCAAACCTGAAGCGCTTCCGTGAACACCCCGCTCTCCGTAGGGGGAGGGGAG
CAGCGGGATCTTGGGTCCCTCATAAGCACTTTGGTTTTACCGCTGCAACCTCACTGTGCCCGCCCCGACCATGCCCTAGCC
CCAGGCTTAGCCGGGCCCATTCAGAGGGGCGAGCACTTGGGGGCACTTCCGGCACTTGGGTGGGACCAAGGAGATGCCACCATAG
ACCTTTCCCTCGCCTTCTTCTCTCTAGTCCGGGTTCAATTTTTCACAGCACCCATCGCCCAAGGGGTACCGAGGGGGGCA
AGGGGTGTCCAGTCCAAGCCACCCCGCTCGCTTCCGCAAACTGTGAGCAAAAAGCAATAGAAGCTCGCCCCCGCCCTT
45 CCCCTTCGAGGATTCGCCGAGTCTGTAGCCTCCCCGATTCAAGTTCCTAGACCTCATGGCTGTCCCTCCACACAGTACCTC
CACTGCACAACCTCGGGCGGGGTGTGACACCTCTCCCCACCCCGACTCCGTGGTTTCCGTATCGTCAACCTTCAGCCGCGG
ACCCGGGAGGGGTCTGGCTTACCTGCTTCTCCCTTCCCATCACTCTTCTGCTTGACAATGTAGCAACCCAGGCCCCCAC
CCAGGCCCTCCCCCTTTTCTCTCTCCCTGACAATAAAGTCTGAATTTGTTCTGCCCTCCG

MYKSLLDKAGLSITSVRFLGDQQSVFVSKDLLKPIQDVNSLRSLTDNQTVSKEFQALIVKHLDESHLLQGDKNLVGSEVRIY
SLDPSTQWFSATVVHGNPSSKTLQVNCEBIPALKIVDPALIHVEVVHDFVTCGNSTRIGAVKRKSSENNSSSVSKQAKSCSEV
SPSMCPVQSVPTTVCKEILLGCTAATPSSNRQQNTPQAANSPPNIGAKLPQGGCHKQSLPEEISSCLNTKSEVLRTKPDVCKAGL
LSSKSSQVGAGDLKILSEPKGSCIQPKTNTDQESRLESTPQPVTGLTKECLVTKTSSKAELDNATAPELQKRLHTASTPDGLS
5 DKPEVEAGVTRLNSCSEKKVGPSDLGSQSNLKETSVKVDHDSCTRSSNKTQTPPARKSVLTDPKLKKLQQSGEAFVQDDSC
VNIVAQLPKCRECLDSLRLKDKDQKDSPVFCRFFHFRRLQFNKHGVLVVEGFLTPNKYDSEAIGLWLPITKNVVGTDLDTAKY
ILANIGDHFCQMVISEKEAMSTIEPHRQVAWKRAVKGVREMCVCDTTIFNLHWVCPRCGFGVCVDCYRLKRKNCCQGAAYKTF
SWIRC VKSQIHEPENLMPQIIPGKALYDVGDIVHSVRKWKIGKANCPCSNRQFKLFSKPAKEDLKQASLSGEKPSLGT MVQQ
SSPVLEPAAVCGEAPSKPASNVK PICPANTSPLNWLADLTSGNVNKENKEKQLTMPILKNEIKCLPPLPLNKSSTVLHTFNST
10 ILTPVSNNSNGFLRNLNLSSTGKTENGLKNTPKILDDIFASLVQNKTTSDLSKRPOGLTIKPSILGFDTPHYWLCNRLCLQD
PNNKSNWNVFRECWKQGPVMVSGVHHKLNSELWKPESEFRKEFGEQEVLDVNCRTNEIITGATVGDVFDGFEDEVNRLKNEKEP
MVLKLDKDFPGEFDRDMMPSRFDDLMANIPLPEYTRRDGKLNLSRLPNYFVRPDLGPKMYNAYGLITPEDRKVGTTLNHLDVD
DAANVMVYVGI PKGQCEQE EEEVLKTIQDGDSEDLTIKRFIEGKEKPGALWHIYAAKDTEKIREFLKKVSEEQGQENPADHDP IH
DQSWYLDRLSLRKLRLHQEYGVQGWAI VQFLGDVFI PAGAPHQVHNLYSCIKVAEDFVSP EHVKHCFWLTQEFRLYSQHTNHD
15 KLQVKNVIIYHAVKDAVAMLKAVNPFVLANLNLPAHWK
20
GGAGAGCATTTTTAGTAGAACATAATTTGGTTTTAGCTGAACGAAAGTCACCTGAAATTTCTGAACGAATTGTACAGTGGCCTG
CAATAACGTACAAACCTCTGTTGGACAAAGCTGGTTTGGGATCCATAACTTCTGTTCTGCTTTCTGGGAGATCAACAAAGAGTAT
TTCCTTCTAAAGACCTTTTGAAGCCTATACAGGATGTAAACAGTCTTCGACTTTCTCTTACGGATAATCAGATTGTCTAGTAAAG
20 AATTTCAAGCTTTGATTGTGAAGCATTAGATGAAAGCCATCTTTTAAAGGTGACAAAACTTAGTTGGTTGAGAAGTAAAAA
TTTATAGCTTGGACCCATCTACTCAGTGGTTTTAGCAACCGTTGTAAATGGAAACCCAGCATCAAAACTCTTCAAGTCAACT
GTGAGGAGATTCCAGCACTGAAAAATTGTGATCCGCTCACTGATTTCATGTTGAAGTTGTACACGATAACCTGTGACATGTGGTA
ATTTGCAAGAATTGGAGCTGTAAACGCAAGTCTTCTGAGAATAATGGAAACCTGGTTTCCAAACAAGCAAAATCTTGCTCTG
AGGCCTCTCCAGTATGTGTCTGTGACGTCTGTACCTACAACAGTTTTAAAGGAGATACTGCTTGGCTGTACTGCAGCAACTC
25 CACCTAGTAAGGCCCAAGACAGCAAGTACTCCCCAGGCTGCCAACTCTCCACCTAACCTTGGAGCAAAATTCCTCAAGGAT
GTCATAACAAAGTTTACCAGAGGAAATTTCTTCTGTCTAAATACAAAGTCTGAAGCTCTGAGAACAAAACAGATGTCTGCA
AAGCAGGGTTGCTCTCAAAGTCTCTCAGATTGGAACCTGGAGACTTGAAATTTCTGACTGAGCCAAAAGGCAGCTGTACTCAGC
CTAAGACAACACATGATCAGGAAACAGATTGGAGTCTGTTCCACAAGCATTGACTGGCCTTCCTAAGGAGTGCCTTACCTACAA
AGGCTTCTTCTAAGGCAGAAATTGGAAATTCGCAATCCTCTGAACTGCAGAAAGCACCTAGAACATGCACCTTCCCCATCGGATG
30 TTTCAAAATCCACCAGAAGTGAAAGCAGGTGTCAATAGTAGTAGCCCTAATAACTGTTTACGGAAGAAAGGTAGAACCCTTACGCTT
TAGCTTGCCGATCACAGAATTAAAGGAATCTTCAGTAAAGTAGATAATGAAAGCTGTTGTTCAAGAAGCAACAATAAAATCC
AGAATGCCCATCCAGGAAGTCGGTTTTGACAGACCCAGCTAAACTCAAAAAGCTGCAACAGAGTGGCGAGGCCCTTCGTACAGG
ATGATTTCTGTGTGAACATCGTGGCACAGTTGCCATAATGCCGAGAGTGTGCTTGGACAGTCTCCGCAAGGATAAGGAGCAAC
AGAAGGACTCACCTGTGTTTTGCCGCTTCTTTCACTTCAGGAGGTTACAATTCAACAAACATGGTGTGTTGCCGGTAGAAGGCT
35 TCTTAACACCAACAAGTATGACAATGAAGCAATTGGCTTTGTGGTTACCTTTAACCACAAACGTTGTGGGGATTGATTGGACA
CAGCAAAGTACATCTTGGCCAACATTTGGAGACCCTTCTGTCAAATGGTGATTTCTGAAAAGGAAGCTATGTCAACTATTGAGC
CACACAGACAGGTTGCTTGAAGCGAGCTGTCAAAGGTGTTGAGAAATGTGTGATGTGTGCGACACCACCATCTTCAACCTGC
ACTGGGTGTGCTCCTCGGTGTGGGTTTGGAGTATGTGTGACTGCTACCGGATGAAGAGAAAGAATTGCCAACAGGGTGTGCTTT
ACAAGACTTTCTCTTGGCTAAAAATGTGTGAAGAGTCAGATACATGAACCAGAGAACTTAATGCCACACAGATCATTCCTGGAA
40 AAGCACTCTATGATGTTGGAGACATTGTTCACTTCTGTAAGAGCGAAATGGGGAATAAAGGCAAACTGCCCTTGTTCAAACAGGC
AATTCAAACTCTTTTCAAAGCCAGCCTCAAAGGAAGACCTAAACAGACTTCTTTAGCTGGAGAAAAACCGACTCTTGGTGCAG
TGCTCCAGCAGAATCCCTCAGTGTGGAGCCAGCAGCTGTGGGTGGGGAAGCAGCTTCAAGCCAGCCGCGCAGCATGAAGCCTG
CCTGTCCAGCCAGCACATCTCCTCTAAACTGGCTGGCCGACCTAACAGCGGGAATGTCAACAAGGAAAACAAGGAAAAACAAC
CAACAATGCCAATTTTAAAGAATGAAATCAAATGCCCTTCCACCCCTCCACCTTTAAGCAAAATCCAGCACAGTCTCTCATACGT
45 TTAACAGACACAATTTGACACCCGTAAGCAACAACAATCTGGTTTCTTCCGGAATCTCTTGAATTTCTTACAGGAAAGACAG
AAAATGGACTCAAGAAATACACCAAAAATCTTTGATGACATCTTTGCCCTCTTGGTGCAAAAATAAGACGACTTCTGATTTATCTA
AGAGGCCTCAAGGATAACCATCAAGCCCCAGCAATTTGGGCTTTTGACACTCTCTCACTATTTGGCTTTGTGATTAATCGCTGTG
GCTTGCAAGACCCCAACAATAAGAGCAACTGGAATGTGTTAGGGAGTGTGGAACAAGGGCAGCCAGTGTGTTGCTGGAG
TGCATCATAAATTGAACCTCTGAACCTTTGGAACCTGAATCCTTCAGGAAAGAGTTTGGTGAGCAGGAAGTAGACCTAGTTAATT
50 GTAGGACCAATGAAATCATCACAGGAGCCACAGTAGGAGACTTCTGGGATGGATTGAAGATGTTCCAAATCGTTTGA AAAATG

AAAAAGAACCAATGGTGTGAACTTAAGGACTGGCCACCAGGAGAAGATTTTAGAGATATGATGCCTTCCAGGTTTGATGATC
TGATGGCCAACATTTCCACTGCCCCAGTACACAAGGCGAGATGGCAAACCTGAATTTGGCCTCTAGGCTGCCAACTACTTTGTTC
GGCCAGATCTGGGCCCCAAGATGTATAATGCTTATGGATTAATCACTCCTGAAGATCGGAAATATGGAACAACAAATCTTCACT
TAGATGTATCTGATGCAGCTAATGTCATGGTCTATGTGGGAATTTCCCAAAGGACAGTGTGAGCAAGAAGAAGTCCCTTAAGA
5 CCATCCAAGATGGAGATTTCTGACGAACCTACAATAAAGCGATTTATTGAAGGAAAAGAGAAGCCAGGAGCACGTGGGCACATAT
ATGCTGCAAAGGACACGGAGAAGATAAGGGAATTTCTTAAAAAGGTATCAGAAGAGCAAGGTCAAGAAAACCCAGCAGACCACG
ATCCTATTCATGATCAAAGCTGGTATTTAGACCGATCATTAAGAAAACGTCTTCATCAAGAGTATGGAGTTCAAGGCTGGGCTA
TTGTACAGTTTCTTGGGGATGTGGTGTATCCCGGCAGGAGCTCCACATCAGGTTTCATAACTTATATAGCTGCATCAAAGTGG
CTGAAGATTTGTCTTCCAGAGCATGTTAAACACTGCTTCTGGCTTACTCAGGAATTCGGATATCTGTACAGACTCATACCA
10 ATCACGAAGATAAATTACAGGTGAAGAATGTTATCTACCATGCAAGTGAAGAATGCAGTTGCTATGCTGAAAGCCAGTGAATCCA
GTTTTGGCAAACCTTAATCTCCCTGCACATTGGAATGAATTACAGGACAGCTGTTCAAACCTTTCAGGCAGGATTCCTGTGGAC
TTTGAGATTCATGTTACCTCATCTTCTTTTTAACTGTACCAACTTGTGAGGGTACTCTGTCTAATGTATATTTCTAGTGT
TACAGACAGTAAATGTGTATATGTAGTAACATTTACAGAACATGCATCCTTAACTGTGACTTCTCACCTAGTGCAGAACTTT
TACCAGGCTGTAAAAGCAAACCTCGTATCAGCTCTGGAACAATACCTGCAGTTATTCTTCAGCTGTTTGGACAACCTTAGATTG
15 GGTTTATAACTATTAGGAATCACTGCACAGTTTATTTGGGTGTGTTTTGTGTCTGAGTCCCTCCCTCATCCCTTAGGCTCCA
GAAGAGCAATGGAGGAAGTGACAGCTAATGTTGCAGTTCTTATGTATGGCATAGGACTGGCATATATAGCAGAAATCAACATA
CTGTACAATTTCTTGGGGTTAACCATCTTTAGTTAAATGGAATTTTAATTTAAATGACGCTTTGCTAATTTTAAGTGTAAAGCA
TTTTGCATTAAAAATATTCATATAAT

21

20 MKAGGGMAGSVSGHSESLASLSRSPQTKAGQKLQHKNPVCVLYIFITKFNKQKPLLNQVILSSVKRN

22

CAGAAGACTACATTAGTGAGATGTAAGAATTATTAAATATTTCCATTTCCGCTTTGGCTACAATTATGAAGAAGTTGAAGGTACT
TCTTTTAGACCACCAGTAAATAATCTCCTTCAAAAAATAAAAAAGAAAAATCATTCAGGAAGAAATGACCTGT
CTAAAAAACCTAAGGAAGATAAATAATAAGAAAGGAAATTTAAAAACATTCCACAAGAAGAAAAATTTATGTTTATACTCC
25 TACTTATGGTTATATCTTATATCTCTATTCAGTGACCTGTCTTTTAAAAAGGCAGTGCTGTCTTACCTCTTGCTAGTGGGTT
AAATGTTTTCAAAATATAGCAGTAGTAGAAGTTTGTATAAAAATTTGTCCTTATTGTTAATTTGTATATAAATGTTAATTAT
TTGATACGAATGTTATGCATTTAGTATGCACATTGAAGTCTAACTGTAGAAGAGTCTAAAAACAAGTTCTCTTTTGCAGATT
ACATACTAATGTTTAAATCTGTGCTCTGTTTAAAGTACTATTATAACTAGAGTAGATCTGAATGAGGATAACCCTAAAAATCAT
GAGGAATGGAAGAAATGGACCTTGAACCTACCTAGGCTTTTATGCATGGCACCTCTTATAATGAAGACACTTTTAAAGTTTTT
30 GTTTTGTGTTTCAATTACCGCTAGATTTTTTTTCTCTTTTAAAAATCCATTTTACTGGAAGTTGGCCAGCAGAGGGAGTAG
AAATTTATTAATTTCTAGTGTGTTGGATTGGGCCCCCTCTTAACAGTACATACTCATTTCCCAAAGCAATCCAAAAACAAATGTG
AACCATTTGGGTTTCAAAATGTTAAGAACCTAAATAGCATGATTTAAAAAATGAAAAATGCTAACACCCAAGAAAAAGAGATAT
TAAGTGCCTTTAAACAACTCCTAGAGTACAAAATGAGTACATCATAATGCTGGCTCTTCTACTAATGAACCATCGAGTGATATT
GAATAAATTTATTTATCTTCTCAGTTTCTTATCTGTAATTTACAATATTAGACTAAGTAAGTTTTCCTCACTTCACTACCAA
35 TTACCTTAGGCTTTTATAATGCTCCGCCCTACTTCAGTCCCATGTTTCAGAAGCTTTTGTCTATTTTTTAACTCATTGATTAAA
TAATGATTAAATGCATTTCTCCACATTTTAATATTGCAAAGGCCCATTTGGAGTTTCTGAAGTGGCTCCACAGAATTGAAATAATTT
CAAATAACTGTAAAGGAACGAAATCTTCACAGAGATGAAGTGGGTTTCCATTAGGTGCTTTGAAATTTGATAACAAATCAT
CAACTTCCACTGGTCAATATATAGATTTTGGGTGCTGAGGCCCAAGATTAGATGCCACTAATCTCCAAAGATTCCCTCCAAT
TATGAAATATTTTAAATGTCTACTTTTAGAGAGCACTAGCCAGTATATGACCATGTGATTAAATTTCTTTTACACTAGATAAAAT
40 TACCTGGTTCAAAAGTGGTTTTTGTATTAAATTTGGTAAATAATATATATAATACACAGACAGGATAGTTTTTATGTGTAAG
TTTTTGGCCAGCTTTAGTTTGGAGACTCCTTGATAAGCTTGCTAACTTTTCAGAGTGCCCTGAGACACTTCCAGCCATCCCTCC
TCCTGCCTTCATTGGGGCAGACTTGCAATGTCAGTCTGACAGTAATTTTTTTTCTGATTGAGAATTATGTAAATTCATACAATG
TCAGTTTTTAAAGTCAAAGTTAGATCAAGAGAATATTTTCAGAGTTTGGTTTACACATCAAGAAACAGACACACATACCTAGG
AAAGATTTACACAATAGATAATCATCTTAATGTGAAAGATATTTGAAGTATTAATTTTAATATATTAATATGATTCTGTATTAT
45 AGTCTTCTGTATGGAATTTTGTCACTTAAGATGAGCTGCAATAAATAATACCTTCAATGGAAAAAAGAAAAAAGAAAAA
AAAAA

23

MGNFRGHALPGTFFFIIIGLWWCTKSILKYICKKQKRTCYLGSKTLFYRLEILEGITIVGMALTMGAGEQFIPGGPHLMLYDYKQ
GHWNQLLGWHHFTMYFFGLLVADILCFTISSLPVSLTKLMLSNALEAFIFYNHTHGREMLDIFVHQLLVVFLTGLVAF

LEFLVRNNVLELLRSSLILLQGSWFFQIGFVLYPPSGGPAWDLMDHENILFLTICFCWHYAVTIVIGMNYAFITWLKSRK
RLCSSEVGLLKNAEREQESEEM

24

5 GAAGCACATCTGGACAGCTGTGCGGCCCTCTTGCGGGCCGACGTCAGCCGAGCAGCTCCCCACGTCTCTCTCTCTCGCCACT
TATTTATTTATTCGTTTTCCTCCAAAGAAGCGACTAGGGACCCAAGTTTAAAAATTCCTCCCCCCTCAATGCGAGACGTGGCCAG
ATCCCATCCAACACACGGTTTAATTTTCATGGGGCTCTGGGATCAAAAGAACAGAAACAGCAACAACAAAGCCAGCCGCTGT
CTGATTTTAAGCTGGCAAAGTGGGAAAAATAAGTGTGAGTAAACAGACCAAGTTGGATCATGGGGAATTTAGAGGTCATGC
CCTCCCTGGAACCTTCTTTTATTTATTTGGTCTTTGGTGGTGTACAAAGAGTATTTCTGAAGTATATCTGCAAAAAGCAAAAGCG
AACCTGCATCTTTGGTTCCAAACATTATTTCTATCGATTGGAATTTTGGAGGGAATTACAATAGTTGGCATGGCTTTAACTGG
10 CATGGCTGGGGAGCAGTTTATTCCTGGAGGGCCCCATCTGATGTTATATGACTATAAACAGGTCACCTGGAATCAACTCCTGGG
CTGGCATCATTTACCATTGATTTCTTCTTTGGGCTGTTGGGTGGCAGATATCTTATGTTTACCATCAGTTTCACTTCTGT
GTCCTTAACCAAGTTAATGTTGTCAAATGCCCTTATTTGTGGAGGCCCTTATCTTCTACAACCACACTCATGGCCGGGAAATGCT
GGACATCTTTGTGCACACAGCTGCTGGTTTGGTGGTCTTTCTGACAGGCCCTCGTTGCCCTTCTAGAGTTCTTGTTCGGAACAA
TGTAATCTTGGAGCTATTGCGGTCAGTCTCATTCTGCTTCAGGGAGCTGGTTCTTTAGATTGGATTGTTCTGTAATCCCCC
15 CAGTGGAGGTCCTGCATGGGATCTGATGGATCATGAAAAATATTTGTCTTCTACCATATGCTTTTGTGGCAATATGCAAGTAA
CATTTGTATCGTTGGAATGAATTATGCTTTTATTTACCTGGTGGTTAAATCTAGACTTAAGAGGCTCTGCTCTCAGAAGTTGG
ACTTCTGAAAAATGCTGAACGAGAAACAAGAAATCAGAAGAAGAAATGTGACTTTGATGAGCTTCCAGTTTCTAGATAAACCTT
TTCTTTTATCATTTGTTCTTGGTTTGTCTTCTGATCTTTTGTTTGGAGAACAGCTGGCTAAGGATGACTCTAAGTGTACTGTT
TGCATTTCCAAATTTGGTTAAAGTATTTGAATTTAAATATTTTCTTTTGTAGCTTTGAAAAATATTTGGGTGATCTTTTCAATTTG
20 CACATCATGCACATCATGGTATTCAGGGGCTAGAGTGATTTTTCAGATTATCTAAAGTTGGATGCCACACTATGAAGAA
ATATTTGTTTATTTGCTTATAGATATGCTCAAGGTTACTGGGCTTGCTACTATTTGTAACCTCTGACCATGGAATTATACT
TGTTTATCTTGTGCTGCAATGAGAAATAAATGAATGTATGTTTGGTGC

25

25 MPSSLWDRFSSSSSTSSPSLPRTPTPDRPPRSWGSATREEGFDRSTSLESSDCESLDSNSGFGPEEDTAYLDGVSLPDFELL
SDPEDEHLCANLMLQLQESLAQARLGSRRPARLLMPSQLVSVQVKELLRLAYSEPCGLRGALLDVCVEQKKSCHSVGQLALDPS
LVPTFQLTLVLRDLRLWPKIQGLFSSANSFPLPGFSQSLTLSTGFRVIKKKLYSSEQLLIEEC

26

30 GCAGCAGGCCAAGGGGAGGTGCGAGCGTGACCTGGGACGGGTCTGGGCGGCTCTCGGTGGTTGGCACGGGTTGCGACACCCA
TTCAAGCGCAGGACGCACCTGTCTTAGCAGTTCTCGCTGACCGCGCTAGCTGCGGCTTCTACGCTCCGGCAGCTCTGAGTTTAT
CAGCAAACGCCCTGGCGTCTGTCTCACCATGCTTAGCCTTTGGGACCGCTTCTCGTCTGCTCCACCTCTCTTCGCCCTCGT
CCTTGCCCCGAACCTCCACCCAGATCGGCGCGCGCTCAGCCTGGGGGTGCGCGACCCGGGAGGAGGGTTGACCGCTCCA
CGAGCCTGGAGAGCTCGGACTGCGAGTCCCTGGACAGCAGCAACAGTGGCTTCGGGCGGAGGAAGACACGGCTTACCTGGATG
GGGTGTCGTTGCCGACTTCGAGCTGCTCAGTGACCTGAGGATGAACACTTGTGTGCCAACCTGATGCAGCTGCTGCAGGAGA
GCCTGGCCAGGCGCGGCTGGGCTCTCGACGCCCTGCGCGCTGCTGATGCCTAGCCAGTTGGTAAGCCAGGTGGGCAAGAAC
35 TACTGCGCCTGGCCTACAGCGAGCCGTGCGGCCCTGCGGGGGCGCTGCTGGACGCTCTGCGTGGAGCAGGCAAGAGCTGCCACA
GCGTGGGCCAGCTGGCAGCTCGACCCAGCCTGGTGCCACCTTCCAGCTGACCCCTGCTGCTGCGCTGGACTCAGACTCTGGC
CCAAGATCCAGGGGCTGTTTAGCTCCGCCAACTCTCCCTTCTCTCTGCTTCAGCCAGTCCCTGACGCTGAGCACTGGCTTCC
GAGTCATCAAGAAGAGCTGTACAGCTCGGAACAGCTGCTCATTGAGGAGTGTGAACCTCAACCTGAGGGGGCCGACAGTGCC
CTCCAAGACAGAGACGACTGAACCTTTGGGGTGGAGACTAGAGGCGAGGAGCTGAGGGACTGATTCCAGTGGTTGAAAACTGAG
40 GCAGCCACCTAAGGTGGAGGTGGGGGAATAGTGTTCACAGGAAGCTCATTGAGTTGTGTGCGGGTGGCTGTGATTGGGGACA
CATACCCCTCAGTACTGTAGCATGGAACAAAGGCTTAGGGGCCAAAGGCTTCCAGCTGGATGTGTGTAGCATGTACCTTA
TTATTTTGTACTGACAGTTAACAGTGGTGTGACATCCAGAGAGCAGCTGGGCTGCTCCCGCCCCAGCCTGGCCAGGGTGAA
GGAAGAGGCACGTGCTCTCAGAGCAGCCGAGGGAGGGGGAGGTCGGAGGTCGTGGAGGTGGTTGTGTATCTTACTGGTCT
GAAGGGACCAAGTGTGTTTGTGTTGTTTGTATCTGTTTCTGATCGGAGCATCACTACTGACCTGTTGTAGGACGCTAT
45 CTTACAGCGCATGAATGTAAGAGTAGGAAGGGGTGGGTGTCAGGATCACTTGGGATCTTTGACACTTGAATAATACACCTG
GCAGCTGCGTTTAAAGCTTCCCCATCGTGTACTGCAGAGTTGAGCTGGCAGGGAGGGGCTGAGAGGTGGGGCTGGAACCC
CTCCCCGGGAGGAGTGCATCTGGGCTTCCATCTAGAAGTGTGTTTACATGAAGATAAGATACTGTTTCTGTTGTAATACACTTG
ATGTTCAAGTATTAAGACCTATGCAATATTTTACTTTTCTAATAAACATGTTTGTAAAAACAAAAA

27

GGPGGPECPMRTPQLALLQVFFLVFPDGVPRQPSSSPSGAVPTSLELQRGTDGGTLQSPSEATATRPVAVPLPTVPTLVTPSA
PGNRTVDLFPVLPIVCDDLTPGACDINCCCDRDCYLLHPRTVFSFCLPGSVRSSSWCVDNSVIFRSNSPFFSRVFMDSNGIRQ
FCVHVNNNLNYFQKLQKVNATNFQALAAEFGGESFTSTFQTQSPPSFYRAGDPILTYFPKWSVISLRLRPAGVGAGGLCAESN
PAGFLESKSTTCTRFKNLASSCTLDSALNAAASYNFVLKVPFSMTDPQNMEFQVPVILTSQANAPLLAGNTCQNVVSQVTYE
5 IETNGTFGIQKVSLSLQGTNLVTEPGASLQQHFILRFRAFQQSTAASLTSRSGNPGYIVGKPLLALTDDISYSMTLLQSQGNG
SCSVKRHEVQFGVNAISGCKLRLLKADCShLQQEIYQTLHGPRPRPEYVAIFGNADPAQKGGWTRILNRHCSISAINCTSCCLIP
VSLEIQVLWAYVGLLSNPQAQVSGVRFLYQCQSIQDSQQVTEVSLTTLNVFVDITQKPQPPRGQPKMDWKWPFDFPFKVAFSR
GVFSOKCSVSPILILCLLLGLVNLFTM

28

10 GGGGGACCCGGAGGCCCTGAATGCCCATGCGCACCCACAGCTCGCGCTCTGCAAGTGTTCTTCTTGGTGTPTCCCGATGG
CGTCCGGCCTCAGCCCTCTTCTCCCATCAGGGGAGTGCCACGCTCTTTGGAGCTGCAGCGAGGGACGGATGGCGAACCCCT
CCAGTCCCTTTCAGAGGCGACTGCAACTCGCCCGGCGTGCTGGACTCCCTACAGTGGTCCCTACTCTCGTGACTCCCTCGGC
CCCTGGGAATAGGACTGTGGACCTCTTCCAGTCTTACCGATCTGTGCTGTGACTTGACTCCTGGAGCCTGCGATATAAATTG
CTGCTGCGACAGGAGTGTCTATCTTCTCCATCCGAGGACAGTTTCTCTCTTCTGCCTTCCAGGACGCGTAAGGTCTTCAAGCTG
15 GGTGTGTGTAGACAACTCTGTTATCTTCCAGGAGTAATTCCTCCGTTTCTTCAAGAGTTTTTCATGGATTCTAATGGAATCAGGCA
GTTTTGTGTTCATGTGAACAACTCAAACCTAAACTATTTCAGAGCTTCAAAGGTCAATGCAACCAACTTCCAGGCCCTGGC
TGACAGAGTTTGGAGGCGAATCATTCACITCAACATTCCAAACCTCAATCACCACCATCTTTTACAGGGCTGGGGACCCCATCTCT
TACTTACTTCCCCAAGTGGTCTGTAATAAGCTTGCTGAGACAACCTGCAGGAGTTGGAGCTGGGGGACTCTGTGCTGAAAGCAA
TCCTGCAGGTTTCTTAGAGAGTAAAAGTACAACCTGCACCTCGTTTTCACAAGAACCTGGCTAGTAGCTGTACCTTGGATTTCAGC
20 CCTCAATGCTGCCTCTTACTATAACTTCACAGTCTTAAAGGTTCCAAGAAGCATGACTGATCCACAGAATATGGAGTTCCAGGT
TCCTGTAATACTTACCTCACAGGCTAATGCTCCTCTGTTGGCTGGAAACACTTGTGAGAATGTAGTTTCTCAGGTCACCTATGA
GATAGAGACCAATGGGACTTTTGGAAATCCAGAAAGTTTCTGTGACTTTGGGACAAACCACTGACTGTTGAGCCAGGCGCTTC
CTTACAGCAACACTTTCATCTTCCGCTTCAGGCTTTTCAACAGAGCAGCTGCTTCTCACCAGCTCTAGAAAGTGGGAATCC
TGGCTATATAGTTGGGAAGCCACTCTTGGCTGTGACTGATGATATAAGTTACTCAATGACCCCTTTACAGAGCCAGGGTAATGG
25 AAGTTGCTCTGTTAAAAGACATGAAGTGCAGTTTGGAGTGAATGCAATACTGGATGCAAGCTCAGGTTGAAGAAGGCAGACTG
CAGCCACTTGACAGAGAGATTTATCAGACTCTTCATGGAAGGCCAGACCAGAGTATGTTGCCATCTTTGGTAATGCTGACCC
AGCCAGAAAGGAGGTTGGACCAGGATCTCAACAGGCACTGCAGCATTTTCAGCTATAAACTGTACTTCTGCTGTCTCATACC
AGTTTCCCTGGAGATCCAGGTATGTGGGCATATGTAGGTCTCCTGTCCAACCCGCAAGCTCAAGTATCAGGAGTTTCGATTCCT
ATACCAGTGCCAGTCTATACAGGATTCTCAGCAAGTTACAGAAGTATCTTTGACAACCTCTGTGAACTTTGTGGACATTACCCA
30 GAAGCCACAGCCTCCAAGGGGCCAACCCAAATGGACTGGAATGGCCATTTCGACTTCTTCCCTTCAAAGTGGCATTTCAGCAG
AGGAGTATTCTCTCAAAATGCTCAGTCTCTCCATCCCTTATCTGTGCTCTTACTACTTGGAGTTCTCAACCTAGAGACTAT
GTGAAGAAAAAGAAATAATCAGATTTTCAGTTTTCCCTATGAGAACTCTGAGGCAGCCACTTATCTTGGCTAAATAGAACCTCA
CCTGCTCATGACCAGAGACATTTAGGATAATAGAGGACCTAACTGAAGGAATCCTTGATATAGAAAGGAGTTATTTTAGAAAA
GCAATAAAAAATATTTTATTCATCATAGCTCTCTGCTTTGGGCTCTGCAGGCCACCAGATACACATGAGGCCCTACTTCTCAAG
35 CTGGGAAGGCCAAGAGCCTTCCTTCAGCCTTTCTGGTTATGTTACACCTAGCTGAATGTTTACAAGGTCTGGATCCATCAGCCC
TCAGGCACAGTTGGGCCAAGCAGAAAGAGAGAACTCTTCTGCTGTCACCTTGAATGAATCAGGAATAGCTTCCCTCTGGACT
GTAGAGGAGCTAACTGTTTGGAACAGAAACTGCTGGCTGTTGATTTTGTCTGGTTCCTTTGCCAACATCTGGGCACACCCCTTT
GCCCAGACACGAGTGGGGAAAGCAGTTCTTTCTCCTCAGTTTCCAAGTAAATGGGGAATCCCAGCTTTCTTTTCTACTAGCAA
ATGACCTTACCATTTATTTCTGCTTTTCTTCCGTTCAITGTGAGGAAAAATAAACTGTTTGAGAGCTTTGTTGTACCAAAA
40 AAAAAA

29

45 RSHFICDLDFVVLHYIFIDFVINMVLGKVKSLTISFDCLNDSNPVYSSGDTVSGRVNLEVTGEIRVKSLSKIHARGHAKVRWTE
SRNAGSNTAYTQNYTTEEVEYFNHKDILIGHERDDNSEEGFHTIHSGRHEYAFSELPQTPLATSPBGRHGSVRYVWKAELHRP
WLLPVKLKKEFTVFEHIDINTPSLLSPQAGTKEKTLCCWFCTSGPISLSAKIERKGYTPGESIQIPAEIENCSSRMVVPKAAIY
QTQAFYAKGKMKEVKQLVANLRGESLSSGKTETWNGKLLKIPVSPSILDCSIRVEYSLMVYVDIPGAMDLFNLPLVIGTIP
LHPFGSRSSSVSSQCSMMNWLSLSLPERPEAPPSYAEVUTEQRRNNLAPVSACDDFERALQGPLFAYIQEFRFLPPPLYSEI
DPNPDSOASDDRPCCSMR

30

5

32

[illegible]

CAAATAAGCACGGAGAAGTGTGGCCAAACTTGGTCTCTGACTCTTATGTATTTGTAAGGCTGGGCTTCAAAATCAAAACAAA
AACCCCAAAAACAGCAGGCAAATGCTTTTAACTCTGACACCGTTGCCATAATCCCTGATACTCAAAGTCTAACAAGAAAGAC
ATGGAATAATTAGCAGCCCATTTTCAGAAAGATCAAAATGATCTAGGGTTCTAATTGCTTTTGCATCCTATTCTTACAAAGTGAT
GTCCCAACAGGGAACAGTAGGAGCTGGAGTGGGATCTCCAAGTCCCAGTTTGAGTGTGGGATGTGCTTCCAGCAGTGCCTTCCC
5 TTTATGAAGACATCACATGGCATCCAGGGCCAGGCAGGCAGCTTGAGGTGCCCTTTACGAGAAAACCGAGCTGGGGCTGGGAGA
GGACAGTTATTGACACTGATGTGAATGAAGTGACAAGATGAGAGCAGAATCGTAAGAGCTTTGAATTTGAAGTGAGTTTTTTT
CCCCCATAAGTTATTTATTCCTTTTCTGTGTAAATATATTTATTTTACTGTGGAGCGCTAACATCTGGATCGTAACATGTG
CAGAATGTATGGTAGGAATGTATTTCTTCTGTAGGAATGTAAATCTGTATTTAAAGGGGGTCCAGCCAGGCCCCAGGTCTTCT
CATGTATGACAGTCCGCATTCTTTTACTCTTCTCTAATATGGGTCTATTTGAAATATGCAAAAGGTATGAGGAATGTTTTT
10 AATACCTCCAAATTTTAAAGAAAGCATCAAAGGGTGTATTTTTTAAAGTTTTTTAGTAGCACTTTCTCTGGATGACAGAA
GGGGCAACCACATGGGCACCTTGTTCATACCAAGGGTGAGCAGTGCCAGAGCCTCCTCTGCACCTCTCGAGTGTCTTTACC
AATTGAGCTTTTATCGCCATAGCCCCTTGGAGTGCCCGAGCTGCCCCGAGGTCAATCAAGGAAAATTTCTTAATGAAATAAGC
TCCAAAGAGCCAAAGTATCAACTTACAGATCGTTTTTAAAGCTTAAATTTATGAACCACCTTTGTGGTAAACAATGAATTATGA
ATACCCAGGGCAGCCTTCTTAAATGACAAATGTAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG
15 GAATTGTCTTAAATTGAAAACCTTGCTGTACAAATTTGAGACCTTTATACATTTTCTGAAAACAATGAAAAGAGTATATTTAACCT
TTTTCTGGCTGTAAATGGTTACCTTCCTGTAACTGCCCGACCTGGAGGCATGGAGTTGTGTGCATCTGCTTATGTAGTACAATT
GTTTTCACTGTTTTCTAAGAATGAGTCTGAATGGTTCTTGAATAATAGCCAGGATCAAATGCTATTGCAGACAAAGCCAAATAAAA
AGTTGCACTTCTTTTGGGGATAACAAGTTTGGGAAGAGAAATGCAGGCCATATGTGCGCATGACCGAGATTTTGAAGAAAGATG
TACATAGTGACATGTTTGGTGCATGGTTTGGAGAGGGCTTTTGTCAAAAAGGAGGTATAACCTTTCCCCACAGACCTGAGA
20 GCTGTGCCCTTTCTATGCAATATTACAGACGTTACATCGGAACCCAGATGGCTGTATTCACATGTAGGTTTGGGCTGTAATCTA
AACAATTGGACAGATTAATGTACATGGAATGAGCAGTCTTACTTTTGTAGTTTTATATTATACAATAACAGTTAAAGATG
33
MTSRDNYKAGSREAAAAAATAAAAAAEEPPVSGAKRKYQEDSDPERSDYEEQQLQKEEBEARKVKSGIRQMRLFSQDEC
AKIEARIDEVVSRAEKGLYNEHTVDRAPLRNKYFFGEGTYGAQLQKRGPQERLYPPGDVDEIPEWVHQLVIQKLVEHRVPE
25 GFVNSAVINDYQPGGCIHVDPPIHIFERPIVSFFSDSALCFGCKFQFKPIRVSEPLVSLPVRGVSIVLGSYAADIEITHCI
RPQDIKERRAVIILRKLRLDAPRLETKSLSSSVLPSPYASDRLSGNNRDPALPKPRSHRKADPDAAHPRILEMDKEENRRSVL
LPTHRRRGFSSENYWRKSYESSEDCSEAAGSPARKVKMRRH
34
AAGTGTACTTCTGCTCTAAAAGCTGCGGAATTCCTCGAGCACTGGCGGCGCCAGCGCTACACGGACCTGCGTGAGAAGCTC
30 AAGTCCATGACGTCCCGGACAACATAAGCGGGCAGCGGGAGCGCGCGCGCTGCGCGACCGCGCTAGCCGCCCGCAGCC
GCAGCCGCGCTGCGCGCGGAACCTTACCTGTGTCCGGGGCCAAGCGCAAGTATCAGGAGGACTCGGACCCCGAGCGCAGCGAC
TATGAGGAGCAGCAGCTGCAGAAGGAGGAGGAGCGCGCAAGGTGAAGAGCGGCATCCGCCAGATGCGCCTCTTTCAGCCAGGAC
GAGTGCGCCAAGATCGAGGCCCGCATTGACGAGGTGGTGTCCCGCGCTGAGAAGGGCTGTACAACGAGCACACGGTGGACCGG
GCCCCACTGCGCAACAAGTACTTCTTCGGCGAAGGCTACACTTACGGCGCCAGCTGCAGAAGCGCGGGCCCGCCAGGAGCGC
35 CTCTACCCGCGGGCGACGTGGACGAGATCCCGAGTGGGTGCACCACTGCTGATCCAAAAGCTGGTGGAGCACCGCGTCATC
CCCGAGGGCTTCGTCAACAGCGCGCTCATCAACGACTACAGCCCGCGGCTGCATCGTGTCTACGTTGACCCCATCCACATC
TTCGAGCGCCCATCGTGTCCGTGCTCTTTAGCGACTCTGCGCTGTGCTTCGGCTGCAAGTTCCAGTTCAAGCCTATTTCGG
GTGTCGGAACCAAGTCTTTCCCTGCGGTGCGCAGGGGAAGCGTGACTGTGCTCAGTGGATATGCTGCTGATGAAATCACTCAC
TGCAATACGGCTCAGGACATCAAGGAGCGCGAGCAGTCATCATCTCAGGAAGACAAGATTAGATGCACCCCGGTGGAAACA
40 AAGTCCCTGAGCAGCTCCGTGTTACCAACCAGCTATGCTTCAGATCGCCTGTCAAGAAACAACAGGGACCTGCTCTGAAACCC
AAGCGGTCCACCGCAAAGCAGACCTGATGCTGCCACAGGCCACGGATCTGGAGATGGACAAGGAAGAGAACCAGCGCTCG
GTGCTGTGCCACACACCGCGGAGGGGTAGCTTCAGCTCTGAGAACTACTGGCGCAAGTCATACGAGTCTCAGAGGACTGC
TCTGAGGCAGCAGGAGCCCTGCCGAAAGGTGAAGATGCGGCGGCACTGAGTCTACCCGCGCGCTCCTGGGAACCTGGGCTC
ATCCTTACGTAGTTGCCCTCCTTTTGTTTTGGGGTTTTGTTTTGTTCATTTGGGGGTTTTTGTTTTTTGTTTTTTGTTTTTT
45 TTTGATTCTATATATTTTCTCTGGTTTTGTGTGCTGTAGGGCTGAAGAATAGAATTGGCCAGGACCTAGGTTCTCATATTCT
TGGTATTCCTCTGGATGGAAGGCTGTGGCATCAATAGGGGACAGGCTGTGCTGGAGTGGCCAGTAGAGGTGGTGGAGC
AGAGCAGCCATTTTAAAGTGGGGTGTATCAGGCTGGGTATTATTTAAAGCAACAATAATGTTTTGTTAAGAAAATTTATTTTG
CTTTTCAGTGTAAATCTTCGAGTGTCTTAAACAAAGTTCACTCTTCTGCTCGCCCTTTCCCTCACTGATGTGCTGCTGCTG
GAGGTCTCTGGAGCTCACAGGCTCTGCTGTCTCCACTTCTCAGCTGCCATCCAGCCCTGCAAGCTCATGCAAAACACCTTT
50 TCTTCTCTGCGGCAGAGTTGTTCAAGTTGCTGGGCAGGGGCTTAAACAGTGCCAGCCCTGCCATCCCAAAGCTATTGTGA

AGCCCCCAGGCGTCCACCCACGCCCAGCTAGCCTGCCATGTCCACAGTTCCCTTGGGCTGCTGAGGGGCTAGTGCAGTGGTC
CTGACCTCTCTTATCAAGAGCACACTTCTTTGCTGGTTCCTCTTTTGAAGCATATGCGTGTGATTATTTGGAACAGTTAGACTT
GCCACGTTGGGTCAGTTTGTAGAAATTGTTTCTAGCTAGAGGACTGGTGTCCCTTCCAAGTCTAGCATTTGGGGTATGGAAAATT
GTGTGTGTGTGGTAGGGTTTTGTGTTTTCTTTTGTAGTTTFTTTTCCCCCTTAGTCTCCTGGCTTTTCTCTTCCCTTCCC
5 TTCTCCACTGGCCAGCTTGGGCTCATCTCATGTATCCTTCTAGGAAGGCGCTGCCCATCTTGTCTGCCGCGCAGCATGCA
TCCAAGGCCAGAGCTCAGGCTGCAGACTGGGCTGGTGCCTCTCCGCTTCAGGTATGGGAGTTGGTGAAGGGGCTTTCAAAA
AAAAAAAAAAAAAAAAAAAA
35
36
MAATVNLELDPIFLKALGFLHLSKSDSAEKLKALLDESLARGIDSSYRPSQKDVEPPKISSTKNISIKQEPKISSSLPSGNNNG
10 KVLTTTEKVKKEAEPADKMSDITEGVDPKKPRLEKPEQSSPITVQSSKDLPMADLSSFEETSADDFAMEMGLACVVCROM
MVASGNQLVEQCENLYHRDCHKPQVTDKEANDPRLVWYCARCTQRMKMAQKTQKPPQKPAVVSVPVAVKDLVKKPETK
LKQETTLAFKRTVEKTSVTSIGNSSASVSSSVTSGLTGWAFAAKTSSAGPSTAKLSSTTQNNTPATSSANQKPVGLTGL
ATSSKGGIGSKIGSNNSTTPTVPLKPPPLTLGKTGLSRVSCDNVSKVGLPSPSSLVPGSSQLSGNGNSGTSGPSGTTSTKT
TSESSSPSASLKGPTSQDHSSML
15 36
GTGAGTTGACCAGAGAAGGATGAGAGAATTGCCAAGTTGACTGAACGCCAAGCTGAGAATGGTGATTCTGAAGATAGAATGCG
TTTGCAATGGCTGCTACTGTGAACCTGGAACCTGATCCCATTTTTTTGAAAGCACTAGGTTTCTTGCAATCAAAGAGTAAAGAT
TCTGCTGAAAAGCTAAAAGCACTGCTTGATGAATCTTTGGCTCGGGGCACTGATTCAGTTACCGTCCATCTCAAAGGATGTG
GAGCCACCCAAAATTTCAAGCACAAAAACATTTCCATTAAAGCAAGAGCCAAAATATCATCCAGTCTTCTCTTCTGGTAATAAT
20 AATGGCAAGGTCTCACAACTGAAAAGGTAAAGAAAGGAGCTGAAAAGAGACCTGCTGTATAAATGAAATCAGACATCACTGAA
GGAGTTGATATTCCAAAGAAACCTAGATTGGAGAAACAGAAACACAGTCATCTCCCATTAAGTCCAAAGTAGCAAGGATTTA
CCTATGGCTGACCTTTCCAGTTTGGAGAGACAGTGTGATGATTTGCCATGGAGATGGGATGGGCTGCGTTGTTGTAGG
CAATGATGGTGGCATCTGGCAATCAATTAGTAGAATGTGAGAGTGCATAATCTCTACACCCAGATTGTCAATAACCCAG
GTGACAGACAAGGAAGCGAATGACCTCGCCTGGTGTGTTGTTGCGCCGATGTACAGACAAATGAAAAGAAATGGCTCAAAAA
25 ACTCAGAAACACCCGAGAAACAGCCCTGCAGTTGTTTCTGTAATCCAGCTGTCAAAGATCCATTGGTTAAGAAACAGAA
ACTAAACTGAAACAAGAGACAATTTCTAGCGTTTAAAGAGAACAGAACTCAAGACATCCAGTTATTTTCAGGAAATCTCT
AGTGCCAGCGTTTCTCTCGTCAGTAAGTGGCTTAACTGGATGGGCACTTTTGCAGCCAAAATCTCTCTGCTGGTCTCTCA
ACAGCAAAATGAGTTCAACAACACAAAACAATCTGGGAAACCTGCTACTTCTGTCAGCTAACCAAGAAACCTGTTGGGTTTGTACT
GGTCTGGCAACATCATCCAAAGGTGGAATAGGTTCCAAATAGGTTCCAAATACAGCACTACGCCCACTGTACCTTTAAACCA
30 CCTCCACCTCTAACCTTGGGTAAAATGGCCTTAGTCTGCTAGTTAGTTGTGACAATGTGAGCAAGTAGGTTCTTCTAGTCCA
AGTAGTTTAGTTCCAGGAAGCAGCAGCCAACTAAGTGGGAATGGAAATAGTGGAAATCAGGACCTAGTGAAGTACTACCAGC
AAAATCTTCAAGATCCAGCAGCTCTCCCTCAGCATCCCTTAAAGGCCAACCTCACAAGATCAGCTCAATGCTATGAAGC
GATTACAGATGGTCAAGAAGAAAGCTGCCCAAAAGAACTCAAGAAGTAATGTGGCAAGTAGGTTTTGTATCATATTAGCCT
AAAGATGAAAGGCTTATTATTATGATATAATCTGTAATACACTGTAATTTAATAAAAGTCTTCATAATCAAAAAAAAAAAAAA
35 AA
37
VDSAMELFLDTFKHPSAEQSSHIDVVRFPVYVINEVRVIPPVRAHSSLPDNRAYGETSPHTFQLDLFFNNVSKPSAPVFDRL
LGSLEYDENTSIIFRPNKSVNTDGLVLRGWYNCLTLAIYGSVDRVISHDRDSPPPPPPPPPPPPPQPSLKRNPKHADGEKEDQF
NGSPPRPQPRGPRTPPPPPDDDDDDPVPLPVSGDKBEDAPHREDYFEPISPDNRNSVPQEGQYSDEGEVEEEQEEGEDEDD
40 VDVEEEDEDEDDRTVDSIPEEEEEDEEEGEDEEGEGDDGYEQISSDEDDGIADLERETFKYPNFDVEYTAEDLASVPPMTY
DPYDRELVLPLLYFSCPYKTTFEIEISRMKDQGPDKENSGAIEASVKLTLLDLREDRGAQWVTALEEIPLIIGLSYLQKLN
TKQDSLGLVDWTMQUALNLQVALRQPIALNVRQLKAGTKLVSSLAECGAQGVTLGLQAGVISGLFELLFADHVSSSLKLNAPKA
LDSVISMTEGMEAFRLGRQNEKSGYQKLELILLDQTVRVVTTAGSAILQKCHFYEVLSEIKRLGDHLAERTSSLPNHSEPDHDT
DAGLERTNPEYENEVEASMDMDLLESSNISEGETIERLINLLEEVFHLMETAPHMTIQPVKSFPPTMARITGPPERDDTYPVLFRL
45 YLHSHHFLFLVTLTLLSIPVTSAPGVQLATKDVLFKLAQSQKGLLFFMSEYEATNLLIRALCHFYDQDEEGLQSDGVIDDAFA
LWLQDSTQTLQCTELFQFQRTASEETHDSDLGLTLNLYLITFNPVGRSAVGHVFSLEKNLQSLITLMEYYSKEALGDSKS
KKSVAINYACILILVVVQSSSDVQMLEQHAASLLKLCADENNAKLQELGKWLEPLKNLRFECINPILIEYVKQIDNLMTPPE
GVGLTTALRVLCNVACPPPPVEGQKDLKWNLAIVQLFSAEGMDTIFIRVLQKLSILTQPWRLHVNMGFTLHRVTTISMARCTL
TLKTMLELLRGGSFEFKDMRVPSALVTLHMLLCISPLSGRLDSDEQKIQNDIIDILLTFTQGVNEKLTISEETLANNWLSLM

LKEVLSSILKVPEGFFSGLILLSELLPLPLPMQTTQVIEPHDISVALNTRKLWSMHLHVQAKLLQEIVRSFSGTTCQPIQHMLR
RICVQLCDLASPTALLIMRTVLDLIVEDLQSTSEDKKQYTSQTTRLLALLDALASHKACKLAILHLINGTIKGDERYAEIFQD
LLALVRS PGDSVIRQQCV EYVTSILQSLCDQDIALILPSSSEGSISELEQLSNSLPNKELMTSICDCLLATLANSESSYNCLLT
CVR TMMFLAEHDYGLFHLKSSLRKNSSALHSLLRKRVVSTFSKDTGELASSFLEFMRQILNSDTIGCCGDDNGLMEVEGAHTSRT
5 MSINAAELKQLLQSKEESPENLFLELEKLVLEHSKDDNLDLSLDSVVLKQMLESSGDPLPLSDQDVEPVLAPESLQNLFN
RTAYVLADVMDQLKSMWFTPFQAEIDTDLVLKVDLIELSEKCCSDFDLHSELSRSLSEPSSPGRTKTKGFKLGKHKHET
FITSSGKSEYIEPAKRAHVPPPRGRGRGGFGQIRPHDIFRQRKQNTSRPPSMHVDDFVAAESKEVVPQDGI PPPKRLPKVSQ
KEQEAPVGLRILLEEITMKVVEARAILTEALFHHYDPLVLQVTAQVLGTVLLEVVGDLPLGLVQIAAVEAQEESLLVEAVVE
VVMYAPLHDKNPFNGNLTVEHFTRTIKIRH
10 38
GGTGGACTCGGCGATGGAGCTGTATTATTTTAGATACTTTTAAACACCCGAGCGCTGAGCAAAGTTCTCATATAGATGTGGTTCC
TTTTCCATGTGTGTTTATATCAATGAAGTCCGAGTCATACCCCGAGGAGTAAGAGCCCATAGCAGTCTGCCAGACAATAGAGC
ATATGGAGAGACATCTCCCATACATTCAATTAGACTTATTCTTCAACAATGTAAGCAAACCAAGTGCCCTGTTTTCGATAG
GTTGGGAAGCCTGGAATATGATGAGAATACTTCCATCATCTTTAGACCTAACTCAAAGGTGAATACTGATGGTCTGGTGCTAAG
15 AGGCTGGTATACTGTCTGACACTGGCAATATATGGATCAGTGGATAGAGTGATAAGTCATGACAGAGACTCTCCACCACCACC
ACCTCCACC GCCACCACCCTCCCGACGACAACCAAGTTTGAAGAAGTAATCCAAACATGCTGATGGGGAGAAAGAGATCAGTT
TAATGGAAGCCCTCCAAGACCACAGCCAAGGGGACCAAGAACTCTCCAGGACCCCTCCACCTGATGATGATGAAGATGATCC
TGTGCTCTGCCAGTGTCTGGTGACAAGGAAGAGGATGCTCCTCATAGAGAAGATTACTTTGAGCCCATTTCTCTGATCGGAA
TTCTGTTCCTCCAGGAAGGCAATATTCTGATGAAGGAGAAGTAGAAGAGGAACAACAAGAAGAAGGAGAAGAAGATGAAGATGA
20 TGTGGATGTAGAGGAAGAAGAGGATGAGGATGAGGATGATCGACGAACAGTAGACAGTATTCTGAGGAGGAAGAGGAAGATGA
AGAGGAAGAGGTGAAGAGGATGAAGAAGGTGAAGGGGATGATGGTTATGAACAAATTTCCAGTGATGAAGATGGAATGTCTGA
CTTGGAACTGAAACATTTAAGTATCCAACTTTGATGTTGAATACACTGCTGAAGACTTAGCTTCAGTTCCCTCTATGACATA
TGATCCATATGACAGGGAGCTTGTACCACTCTTATACCTTCAGTTGTCCATACAAGACTACTTTTGAAATTTGAAATCAGTAGAAT
GAAGGATCAAGGTCAGATAAAGAAAATTCAGGGCAATCGAAGCCTCAGTGAAGTTAACAGAACTCTTAGATTGTATAGAGA
25 AGATAGAGGTGCAAAATGGGTAACAGCTTTAGAAGAAATTCAGTTTAAATAAAGGGTTAAGCTATTGCAATTGAAAAA
CACAAAACAAGACTCCCTTGCCAGTTGGTAGACTGGACCATGCAAGCTTTAAATTTACAAGTAGCGCTTCGCCAACCTATCGC
CTTAAATGTTTCAGAGCTCAAAGCTGGGACCAAAATAGTGTCTCACTAGCAGAATGTGGGGCTCAAGGAGTTACAGGACTGCT
ACAAGCAGGAGTGATCAGTGGATTATTGAACTTCTGTTGCTGATCAGTATCATCTCTCTTAAGTTAAATGCTTTTAAAGC
TTTGACAGCTGTCTATTAGTATGACAGAAGGAATGGAAGCTTTTAAAGAGGTAGGCAGAAATGAAAAAGTTATCAAAGCT
30 TCTGGAACCTCATACTTTTAGATCAGACTGTGAGGGTTGTTACTGCTGGTTCAGCTATTCTCCAAAAATGCCATTTCTATGAAGT
CTTGTGAGAGATTTAAAGACTTGGTGACCATTTAGCAGAGAAGACTTCATCTCTTCTTAACCACAGTGAACCTGATCAGCAC
AGATGCTGGACTTGAGAGAACAAACCCAGAATATGAAAATGAGGTGGAAGCTTCTATGGATATGGATCTTTTGGAATCTCAAA
TATAAGTGAAGGGGAAATAGAAAGCTTATTAACCTCTAGAGAAGTTTTCATTTAATGGAAGCTGCCCTCATACAATGAT
CCAACAACCTGTTAAGCTTTTCCAACGATGGCAGCAATTAAGTGGACCTCCAGAGAGGGATGATCCATACCTGTTCTCTTTAG
35 ATATCTTCACAGTCATCACTTCTTGGAGTTGGTTACCTTGCTTCTGTCAATTCCAGTAACAAGTGCTCACCTGGTGTGCTGCA
AGCCACAAAAGATGTTTTGAAGTTCTTGACAGTCACAGAAGGGTCTTCTTTTTTTTATGTCGGAATATGAAGCAACAAATTT
ATTGATCCGAGCTCTGTGCTACTTTTATGATCAAGATGAGGAGGAAGGTCTCCAATCTGATGGTGTATTGATGATGCATTTGC
CTTGTGGCTACAGGACTCAACACAGACATGCAATGTATTACAGAAGCTTTCAGCCATTTTCAGCGTTGTACAGCCAGTGAAGA
AACAGACCAATTCAGATCTCTTGGGAACCTGCAACAATCTTATTGTGATTACTTTAATCCTGTGGGAAGATCAGCTGTTGGCCA
40 GTTTTTTAGTCTGGAGAAAAATCTCCAAAGCTTATTACTCTAATGGAGTACTATTCCAAAGAAGCCTTGGGTGATTCCAAATC
TAAGAAGTCAGTAGCTTATAATTACGCATGCATACTTATTTTGGTGGTGGTTCAGTCTTCCAGTGATGTTCAAAATGCTAGAACA
ACATGCAGCATCTCTCTTGAAGCTTTGTAAAGCAGATGAAAATAATGCTAAATGCAAGAACTTGGCAAGTGGCTTGAACCTCT
GAAAAACCTTAGATTTGAAATTAAGTGCATCCCAAACTTAATGAGTATGTTAAGCAGAAATATCGATAACTTGATGACCCAG
AGGAGTTGGCTTACCAC TGCTTACGTGTTCTCTGTAATGTTGATGCCCCACCACTCCTGTTGAAGGTCAACAGAAAGATCT
45 GAAATGGAATCTTGCCGTTATTAGCTTTTTCTGCTGAAGGAATGGACAGCTTTATTGAGTTCTGCAAAAAATGAACAGTAT
TCTGACTCAGCCTTGGAGGCTCCATGTCAACATGGGGACTACCTTTCACAGAGTTACTACTATTTCATGAGCTCGCTGCACACT
CACTCTTCTTAAACTATGTTAAGCAAGCTCCTGAGAGGTGATCTTTGAGTTAAGGACATGCGGTGTTCTTCAGCGCTTGT
TACTTTACATATGCTCCTGTGCTCTATCCCTCTCAGGTGCTTTGGATAGTGATGAACAGAAAAATTCAGAAATGATATCATTTGA
TATTTTACTGACTTTTACACAAGGAGTTAATGAAAACTCACAATCTCAGAAGAGACTCTGGCCAAATAACTTGGTCTTTAAT
50 GTTAAAAAGAAGTTCTTTCTCAATCTTGAAGGTCTCTGAAGGATTTTTTCTGGACTCATACTCTTTTCAGAGCTGCTGCCTCT
TCCATTGCCCATGCAAACTCAGGTTATTGAGCCACATGATATATCAGTGGCACTCAACACCCGAAAAATTTGTGGAGCATGCA

CCTTCATGTTCAAGCAAAGTTGCTCCAAGAAATAGTTCGCTCTTTCTCTGGCACAACCTGCCAGCCCATTCACATATGTTACG
GCGTATTTGTGTTCAATTGTGTGACCTTGCCCTCACCACCTGCACCTCTGATTATGAGAACTGTGTTGGATTGATTGTAGAAGA
CTTGCAAAGCACTTCAGAAGATAAAGAAAAACAGTATACTAGCCAAACCACCAGGTTGCTTGCTCTTCTTGATGCTCTGGCTTC
ACACAAAGCTTGTAAATTAGCTATTTTGCATCTAATTAATGGAACATTTAAAGGTGATGAAAGATATGCAGAGATATTCCAGGA
5 TCTTTTAGCTTTGGTGGCTCTCCTGGAGACAGTGTATTCGCCAACAGTGTGTGAATATGTCACATCCATTTTGCAGTCTCT
CTGTGATCAGGACATTGCACCTTATCTTACCAAGCTCTTCTGAAGGTTCTATTTCTGAACTGGAGCAGCTCTCCAATTCTCTACC
AAATAAAGAATTGATGACCTCAATCTGTGACTGTCTGTTGGCTACGCTAGCTAACTCTGAGAGCAGTTACAACCTGTTTACTGAC
ATGTGTCAGAACAATGATGTTTCTTGAGAGCATGATTATGGATTATTTTCAATTTAAAAAGTTCTTTAAGGAAAAACAGTAGTGC
10 TCTGCATAGTTTACTGAAACGAGTGGTCAGCACATTTAGTAAGGACACAGGAGAGCTTGCATCTTCATTTTGAATTTATGAG
ACAAATCTTAACTCTGACACAATTGGATGCTGTGGAGATGATAATGGTCTCATGGAAGTAGAGGGAGCTCATACATCACGGAC
GATGAGTATTAATGCTGCAGAGTTAAACAGCTTCTACAAAGCAAAGAAGTCCAGAAAAATTTGTTCTTGAAC TAGAGAA
GCTTGTTTTGGAACTTCAAAGATGATGACAATCTGGATTCTTTGTTGGACAGTGTAGTTGGACTTAAGCAGATGCTGGAGTC
ATCAGGTGACCTTTACCTCTCAGTGACCAGGATGTAGAACCAGTACTTTCAGCTCCAGAATCTCTTCAGAATCTGTTTAAACA
TAGGACTGCCTATGTGCTTGCTGATGTCATGGATGATCAGTTGAAATCTATGTGCTTCACTCCATTTTCAGGCTGAAGAGATAGA
15 TAGAGATCTGGATTGTTGTTAAAGTTGACTTAAATGAACTCTCTGAAAAATGCTGTAGTGACTTTGATTTGCACCTCAGAATTAGA
GCCCTCATTTTGTGAGAACCATCATCTCCAGGAAGAACCAAGACTACTAAAGGATTCAAACCTGGGAAGCACAAGCATGAGAC
CTTTATAACGCTCAAGTGGAAAACTGAATACATTTGAACCTGCCAAAAGAGCTCATGTTGTGCCACCACCAAGAGGAAGGGGCAG
GGGAGGATTTGGACAGGGTATACGACCTCATGATATTTTTCGTGAGAGAAAACAGAACAAGTAGACCACCATCTATGCATGT
GGATGACTTTGTTGCTGCTGAAAGTAAAGAGTGGTTCTCAAGATGGAATACCTCCACCAAAACGGCCACTCAAAGTATCACA
20 GAAGGAACAAGAGGCTCCAGTTGGAGTGTCTAGAATACTCTCGAGGAAATACAATGAAAGTCGTGGAGGCCAGAGCAATTTT
AACAGAGGCCCTCTTCCACCATTTACGACCCCTTAGTTCTACAGGTTACCGCCCAAGTCTCGGGACCGTGCTTCTAGAGGTCGT
GGGGGACTTGGACCTTCTTGGGCTAGTGCAAATAGCGGCAGTGGAGGCTCAAGAGGAAAGTTTGTAGTGGAGGCAGTGGTAGA
GGTCGTATGTACGCTCTTTACACGATAAAAACTCTTTTGGGAACATCTTAACTGTATATGAACATTTACAGGAGACAATAAA
AATAAGACATTGAAGGACCAATTTAGACTTAGCAGTTATCTGGAGACATCTGAGAGAATATTTTATCTGAAGAAAGCAGAATT
25 TGTTTGATACCTAACAAGATTTCAATAAAAAATCCAACTTTGTATGT

39

MAARDSSEEDLVSYGTGLEPLEEGERPKKPIPLQDQTVRDEKGRYKRFHGAFFSGGFSAGYFNTVGSKEGWTPSTFVSSRQNR
DKSVLGPEDFMDDEEDLSEFGIAPKAIVTTDDFASKTKDRIREKARQLAAATAPIPGATLLDDLITPAKLSVGFELLRKMWKEG
QGVGPRVKRRPRRQKPDPGVKIYGCALPPSGSEGEEDDDYLPDNTVPFAPKDVTPVDFTPKDNVHGLAYKGLDPHQALFGTSG
30 EHFNLVSGGSEERAGDLGEIGLNKGRKLGISQAFGVGALBEEDDDIYATETLSKYDTVLKDEEPGDGLYGTWAPRQYKNQKES
KDLRYVKGILDGFLASKPLSSKKIYPPPELPRDYRPVHYFRPMVAATSENSHLLQVLSESAKATPDGTHSKHLNASKRAE
LLGETPIQGSATSVLEFLSQDKERIKEMKQATDLKAAQLKARSLAQNAQSSRAQLSPAAAAGHCSWNMALGGGTATLKASNFK
PFAKDPEQKRYDEFLVHMKQGQKDALERCLDPSMTEWERGRERDEFARAALLYASSHSTLSSRPTHAKEEDSDQVEVPRDQE
NDVGDQKQSAVKMKMFGLTRDTFEWHPDKLLCKRFNVPDPYPDSTLVGLPRVKRDKYSVFNFLTLPETASLPTTQASSEKVSQH
35 RGPDKSRKPSRWDTSKHEKKEDSISEFLSLARSAEPPKQSSPLVNKEEHAPELSANQTVNKDQDAQEGBSRPSMDLFRA
IFASSSDEKSSSEDEQDSEDDQAGSGEANFQSSQDIDLGETSSVAHALVPAPQEPSPFPQKMQIDEREFPRLPPVFCP
NARQTLVPQKEKHKKNKDKHKAKKEHRRKKEKKKKHRKHKGKQKNKKPEKSSSSSDSDSDSQSDEETADVSPQELLRLK
SLPLRRQ

40

GTGTTAGAATTTCTGTCCCAAAAAGACAAAGAGAGAATCAAAGAAATGAAGCAGGCAACTGACCTGAAAGCAGCTCAGCTCAAG
GCCAGGAGTCTGGCCAGAACGCTCAGAGCAGCAGAGCCAGCTCTCCCTGACAGCGGCTGCTGGGTACTGCTCTTGGAAATG
GCATTAGGTGGTGGACGGCCACCTTAAAGCCAGCAACTTCAAGCCTTTTCGCCAAAGATCCGGAAGCAAAAGCGATACGAC
GAGTTCTTAGTACACATGAAAACAGGTCAGAAAGATGCTCTGGAACGCTGTCTGGACCCAGCATGACGGAGTGGGAGCGAGGC
CGTGAGCGGGATGAGTTTGCCCGGGCGGCCCTGCTGTACGCATCTTCCATTTCGACCTTGCTCTCCAGGTTCACTCACGCCAAG
45 GAGGAGGATGACTCAGATCAGGTTGAAGTCCCTCGAGACCAAGAGAATGATGTGCGGGATAAGCAGTCCGCTGTGAAGATGAAG
ATGTTTGGGAAGCTCACCCGAGACACGTTTGAGTGGCACCCCTGACAAGCTTCTATGTAAGAGATTTAATGTCCCTGACCCTTAT
CCAGATTCAACTTTAGTTGGCTTACCAAGAGTGAAGCGTGACAAGTACTCAGTCTTCAACTTTCTGACGCTCCAGAGACAGCT
TCCTTGCCACCACTCAAGCATCAAGTGAAGAAAGTATCACAGCAGCGAGGTCCCGACAAATCAAGAAAAACCATCCAGATGGGAT
ACCTCTAAACACGAAAAGAAAGAGATTCCATTAGTGAATTTTTAAGTTTGGCTAGATCAAAAGCCGAGCCACCTAAACAACAG
50 TCCAGCCCTTAGTAAACAAAGAGGAAGAGCATGCACCAGAATCATCCGCAATCAGACAGTCAACAAAGATGTGGACGCACAG

GCTGAAGGAGAAGGGAGCCGCCATCCATGGACTTATTCAGGGCCATCTTTGCCAGTTCCTCAGATGAAAAGTCTCATCTCC
GAGGATGAGCAAGGTGACAGTGAAGATGATCAGGCAGGCTCTGGGGAGGCCAACTTCCAAAGTCCCAAGACACTGACTTGGGG
GAAACATCATCTGTGGCTCAGCTCTTGTGCCAGCACCCAGAGCCGCCACCTTCCCTTCCCGATACAAAAGATGCAGATAGAT
GAAAGAGAAGAGTTCCGGCCCGGCTGCCTCCCGTCTTCTGCCCAATGCTCGTCAGACACTTGAGGTTCCTCAAAAAGAGAAA
5 CATAAAAAGAACAAGACAAGCACAAGGCCAAGAAAGAGCACAGGCGGAAGAAAGAGAAAGAAAACACAGGAAGCACA
CACAAAGGCAAGCAAAGAATAAAAAACCAGAGAAAAGTAGTAGCTCCGAGAGTTCCGACAGCAGCGACAGCCAGAGTGACGAG
GAAACCGCAGACGTGTCCGCCAGGAGCTGTGAGACGGCTGAAAAGTCTTCCACTAAGAAGGCAGTAATTGAATGCTGCCCTG
GCTCGTCTAGAAATCATTTCTCTCCATGATGGAAGCCAGTGATTTTTCAGTTAACGCATTGTACAGAGTGATTTTATATGTA
AATTCCTGCTGTAAATAATTTTAAACCTTGACATTTCAAAGACTGCCTTGAAAGSTCGCAGAAATGATTTCTATTTTAA
10 TTTTCAGTTAGTATCGGGGAAAAAATCCAGATTGAACAGTTTAATTAAGTGAATTTTCT
41
MAQVSINNDYSEWDLSTDAGERARLLQSPCVDTAPKSEWEASPGGLDRGTTSTLGAIFIVVNACLGAGLLNFPAAFSTAGGVAA
GIALQMGLVFIISGLVILAYCSQASNBRTYQEVVWAVCGKLTGVLCEVAIAVYTFGTICIAFLIIIGDQDKIIAVMAKEPEGA
SGPWYTDKFTISLTAFLFILPLSIPREIGFQKYASFLSVVGTWYVTAIVIIKIYIPDKEMTPGNILTRPASWMAVFNAMPTIC
15 FGFQCHVSSVPFNSMQPEVKTWGGVVTAAVMIALAVMGTGICGFLTFGAAVDPDVLSPSEDMAVAVARAFIILSVLTSY
PILHFCGRAVVEGLWLRVQGVPEEDVGRERRRRVLQTLVWPLLLTLLALFIPDIGKVISVIGGLAACFIFVFPGLCLIQAKLS
EMEEVKPASWWVLVSYGVLVLTLAGAFIFGQTTANAIFVDLLA
42
ATGTGATGCATGCTCACGTGTCTCCGACGCCGCTCGGGAAAGAATCCCCAAGCTCCATTTTCATGAGTAAGCGTGAGAGCCGC
20 TCAGTTTCTCCAGCTCTGCTGAAGCCAGCACAGAAGTAGCCCAAACCTTCCCTCTGCTGACAGCAAATTTAGGCAAAGTCT
TGAGAAAGAAGAAATTGGGTCCAGAAAGGGAAGTGAGGAGAATCAGATCCCAGACCTTTGGGGAGAAGGAGCAACCGCTCTGG
CACAGCCCATCAGGGAGAAAGAGCAGGTTGAGAAGAGTCTTAAGCTAACAGCCCCAACAGGTGGGTGTTGCTCAGCTCCCTGA
GGCATGTGGTTGTAAGGCAGAACCCACAGACCTTGCAGGAAGAAGGCTCTCGGGGCCATGGCCAGGTGAGCATCAACAATGAC
TACAGCGAGTGGGACTTGAGCACGGATGCCGGGAGCGGGCTCGGCTGCTGCAGAGTCCCTGTGTGGACACAGCCCCAAGAGT
25 GAGTGGGAAGCTCTCTGGGGGTCTGGACAGAGGCACCACTTCCACACTTGGGGCCATCTTCATCGTGTCAACCGGTGCTG
GGTGACGGGTACTCAACTTCCAGCAGCTTTCAGCACTCGGGGGGCTGGCAGCAGGCATCGCACTGCAGATGGGTATGCTG
GTTTTCATCATCAGTGGCTTGTCTATCTGAGGTGGCCATCGCTGTCTACACCTTTGGCACCTGCATTGCCCTTCTTAATCATC
GTGTGTGGCAAGCTGACAGGTGTCTATGTGAGGTGGCCATCGCTGTCTACACCTTTGGCACCTGCATTGCCCTTCTTAATCATC
ATTGGCGACCAGCAGGACAAGATTATAGCTGTGATGGCGAAAGAGCCGAGGGGGCCAGCGGCCCTTGGTACACAGACCGCAAG
30 TTCACCATCAGCCTCACTGCCTTCTCTTCATCTGCCCTTCCATCCCCAGGGAGATTGGTTTCCAGAAATATGCCAGCTTC
CTGAGCGTGTGGGTACCTGGTACGTACAGCCATCGTTATCATCAAGTACATCTGGCCAGATAAAGAGATGACCCAGGGAAC
ATCTGACCAGGCCGGCTTCTGGATGGCTGTGTTCAATGCCATGCCACCATCTGCTTCGGATTTCAGTGCCACGTCAGCAGT
GTGCCGCTCTTCAACAGCATGCAGCAGCTGAAGTGAAGACCTGGGGTGGAGTGGTGACAGTGGCATGGTCATAGCCCTCGCT
GTCTACATGGGGACAGGCATCTGTGGCTTCTGACCTTTGGAGCTGCTGTGGATCCTGACGTGCTCTGTCTATCCCTCGGAG
35 GACATGGCCGTGGCCGTGCCCCGAGCCTTCATCATCTGAGCGTGCTACCTCTACCTATCTGCACTTCTGTGGCGGGCG
GTGTGGAAGGCCGTGTGGCTGCGCTACAGGGGGTGCCAGTGGAGGAGACGTGGGGCGGAGCGGGCGGCGAGTGTGCAG
ACGTGCTGTGGTTCTGTCTACCTGTCTGCTGGCGCTTTCATCCCTGACATCGGCAAGGTGATCTCAGTCAATTGGAGGCCTG
GCCGCTGTCTCATCTTCTGCTTCCCAGGGCTGTGCCTCATTCAAGCCAACTCTCTGAGATGGAAGAGGTCAAACAGCCAGC
TGGTGGGTGCTGGTCAGCTACGGAGTCTCTTGGTCACCTTGGGAGCCTTCATCTTCGGCCAGACCACAGCCAAACGCCATCTTT
40 GTGGATCTCTTGGCATAACCACTGCCTCCCAGGGAACACAAGGCCTTTGCCATTGGTCGAGGAACCATCTCTTAGAGCTATG
GGCCATCTTAGTCCACGATCATTTCAACTGGTGGGATGACATCCGACATCTCTTCCAGGGAAGTGGGCAAACTCAGGCCC
CACACCTCTGGACAGCTCAAATCCAGTCCCTTCTGCTCCCGAGTCTGCGAGTGGCGGATGGCGGCAGGAAGTCTCACAT
CATAGAGGACCCCTCTCTCTCTCCAGTTCTCAACTTCTCCATGCTTGAATCCACGGGTGAAGAGAGTCCGTTAGATCTCATAA
45 GAAAGAATCCAGTCTGACTTCCCTCTGGAGAATGACTATGGACAGAAGGCCACCATCTCCACAGAGCACCCTGTCTGAGTAG
GGGTGTGCTCATTTACCCAGGCCAGTGTAGCTTCTCAGGAGCCTGGCCATTTCCAAAGGTAGCAGTGAAGTCTGCTGCTAAT
CATAGTCAGGTAGATTAGACCTTGTCCACACCTTCTTGGGCAACCCCCACCATGAACCTGTGAGCCTCTTTCCCATAGCTAAT
AGACATTTCCAGGCCCTTGGGGGCCCAACCTGTCTCTTTCATCAAACCTGATGGTCCAGGCTGGGCATCCCTCTCTCTCTCC
ATCCCCAGACATCACCAGGTCTAATGTTTACAAACGGTGCCAGCCCGGCTCTGAAGCCAAGGGCCGTCCCGTCCACGGTGTGT
TGAGTATTTCTCTCTTAGCTTTCCCCATAAGGTTGGGAGTATCTGCTTTTGTGTCTGAGATGGGCCCTCTTTTCAGAGGCCGC
50 AGGTGGGTGATGGAGAAGGCTGAGAACCTTTTCAGACCTCTGTGTGGGCTGGGCTGGTCAGAAATCAGGCTGTACCTCCCCGAC

ACCTTCTTTTTCAGTGATGTTTTCTCTTCTCCCTGCCTTTCCTCTGCCTCCTCCCTGCCAGCCCTAGCGTGACTACCCAGAGAC

43

MAQVSINNDYSEWDLSTDAGERARLLQSPCVDTPAKSEWEASPGGLDRGTTSTLGAIFIVVNACLGAGLNFPAAFSTAGGVAA
5 GIALQMGLVFIIISGLVILAYCSQASNERTYQEVVWAVCGKLTGVLCEVAIAVYTFGTICAFLLIIIGDQDKIIAVMAKEPEGA
SGPWYTDKFTTISLTAFLFILPLSIPREIGFQKYASFSLVVGTVVYTAIVIIKIWPKEMTPGNILTRPASWMAVFNAMPTIC
FGFQCHVSSVPVFNMSMQQPEVKTWGGVVTAAMVIALAVMGTGICGFLTFGAAVDPDVLLSYPSEDMAVAVARAFIILSVLTSY
PILHFCGRAVVEGLWLRVYQGVPEEDVGRERRRRVLQTLVWFLTLALLLALFIPDIGKVISVIGGLAACFIFVFPGLCLIQAKLS
EMEEVKPASWWVLVSYGVLVLTGAFIFGQTTANAIFVDLLA

10 44

GAGGCCAAGAGCGCGGGCGGCGAGGCAAGATGGCGGCAACCAAGAGGAAACGGCGTGGAGGCTTTGCAGTTTCAGCGAAGAAGC
CAAAAAGAAACGAAATAGATGCGGAGCCGCCAGCTAAGCGGCACGCCACAGCAGAGGAGGTGGAGGAAGAAGAGAGGGACCGGA
TCCCAGGCCCGCTTTGCAAGGGAAAGTGGAAAAATAAGGAACGATTCTCATCTTTTCTTCCAGAGGAATAAAATTTTAGAACAA
GACATTTAATGCAGGACTTGAGAATGTTGATGCCTCATTTCTAAAGCAGATACTAAAATGGATCGTAAGGATAAGCTATTGTGTA
15 TTAACGAGGTTTGTGAAATGAAGAACTGTAATAAATGCATCTATTTGAAGCTAAGAAAAAACAGGATCTCTATATGTGTAAG
AGAATGTATTAAGATTTTGGTTAAACTCATTTAAGTGGATTGTCTTTGTACCTTTTATGTTATAGACTCCTAGTGTTCATG
TTAAGGGTTGAGAATGAAGCAAACTTTGATTTACGAAACTTGATACCTTTAAAGAAAATTAATGTATAAATATTATAAGCA
TATTATTTATGGTGAATAGGTGGTAAGCATTGACTACATTTGCTAATTAGTGAAAAAGCTTACTCCATATCTTTCAGGCTTTC
AAATTCACCTCACGACCATCTGCTAAATTCCTTGTTCAAAATAGTAAGTTGACTCAATAAATTTTTTTAGATGAGGAAATTA
20 AAGTAATCTTTTGAATAGTTGTGCAAAAGTTACAACATTTTTGTGTTTTGTGTAATTTTCTAGTTCATACCTCGCTGAAC
TGAAGATGACTGGAAACTGTTGAAAGGTTCTCGGCCCTTTTGTCTTTTGACCTGTAAAGTTTCTCATTTCAGTGTATGAGGTC
TAATTTTCTTATCTGGCATGGTTGTTTTAGTAGATATAGTTGTGTTATCAATTTAGTTTCACCCCCACCTTGGTTTTATGGA
TTATCAATTTTTCAGGTACATTTATAATGAGGTTATAGCCAAGTTATAGAAACAAATGAGCAACAGTTTTAAACAAGGCAGT
CTACTGTTAAATAATATGCTTACCTTATTTTATAGGCTTTTGATGAATTACCACATTATGCTTTGTTAAAAGAACTCTTAATT
25 CAGGTAAATATCTTTAAATTAGCTATCCAAAATATACATGATATATCTTGAATAAAGGAACATAACACTTTTTTAAGTAG
ATCTTTTAGTACACCACGGTATCATCCAAAAGCCAACCATTTGTGGACCACGTGTTTACTTTTACCATTTTGGATAATAGGATA
TGGTTTCGGAACCTTCAGATCATAGAAGAAGATGCTGCTCTGTAGAAAATAGGACCTCGTTTTGTCTTAAATCTCATAAAGATT
TTCCAGGGAAGTTTTGGAGGACCACTTTATATGAAAATCCTCACTACCAGTCACCAACATGGTAAGCGGTTGTGTCATTTCAG
TAGCTTTCAATGTACCAGAACCCTTTGGCCAAAATGATCTGTGAAGTAATTGCTGTTTTATTACCTGTGAAACTGAAATTTGGT
30 TTTTCTGCTATAGAAAATTTGAAATTTACTCCCTTTTTCATGACACTGGCTGAAAGGATATATGGTTTCAACTGAAAGTCTTT
GCCTTATATAAAATGTTTATTTTATTTCAAAGCATCTAATCTTTTTTTTTTTTTTTAGCATCGCGGTGTCATAAGATCCAT
CACAGCTGCAAAATACAGAGAGAAACAGCAAGTGAAGATGTGCAAAAACCTGAGAAAGAAAGAGCCGAAGACTCTTCTCCACA
TGATCCCACTGCAGATGTTTTTGTAAACACAGCTGAGGAGAAACCAATAGAAATACAGTGGGTAACCAGAGCAAAAGTTGA
TTTGAAAGCAAGAAAGAAACGGATTACAAAAGGCAAGAAAAATGAACAGAGGATGGACAGTGGGAAAAACAAAATAAGTCAA
35 TGGAAACCTGATTTGTTTTTCAGTTACTTTATATTTATTTGTATTCAATGTGTAAATACTTTTATTATCTAATACTATCTTAC
GTCTAATTAGTGTAGCATTTACAAGAAAGAAAAATTAAGATCTTAAATCAGTGATTATCTTTTCTAAATAAAATATCACCAG
AATTCATCAGTT

45

MDWKEVLRRRLATPNTCPNKKKSEQELKDEMDLFTKYYSEWKGRKNTNEFYKTIIPRFYYRLPAEDEVLLQKLREESRAVFLQ
40 RKSRELLDNEELQNLWFLLDKRQTPPMIGBEAMINYENFLKVGKAGACKQFFTAQVFAKLLHTDSYGRISIMQFFNYVMRKV
WLHQTRIGLSLYDVAGQGYLRESLENYILELIPTLPQLDGLEKSFYSFYVCTAVRKFFFLDPLRTGKIKIQDILACSFLLDL
LELRDEELSKESETNWFSAPSALRVYGYLNLDKDHNGMLSKEELSRHGATMTNVFLDRVFQECLTYDGMEDYKTYLDFVLA
LENRKEPALQYIFKLLDIENKGYLNVFSLNYFFRAIQELMKIHQDPVSFQDVKDEIFDMVKPKDPLKISLQDLINSNQGDV
TTILIDLNGFWTYENREALVANDSENSADLDDT

45 46

AAAAATCCGGCTTGGTGAGCTTGGGTCGCCTCTGAAGGAGAACCATTTTCCATCTCTTTTCATAGTTTTTCCCCCAGTCAGCGT
GGTAGCGGTATTCTCCGGCAGTGACAGTAATTGTTTTGTCTCTTTAGCCAAGACTTCCGCCCTCGATCAAGATGGTGGTTG
GACGGCTTCTTAACCTTTACGGGGCTGGCGGTGCTGACGCTGAGCTGGTAGGGGTGGAGCAGGTAGGAAACAGCAAAATGCA

- GAAGCTGCTGCGCGGGAGTCGGCCATGGACTGGAAGAAGTCTTCGTCGCGCGCTAGCGACGCCAACACCTGTCCAAACAAA
AAAAAAGTGAACAAGAAATTTAAAGATGAAGAAATGGATTTATTTACAAAATATTTACTCCGAATGGAAAGGAGGTAGAAAAAAC
ACAAATGAATTTCTATAAGACCATTCCCCGGTTTTATTATAGGCTGCCTGCTGAAGATGAAGTCTTACTACAGAAATTAAGAGAG
5 GAATCAAGAGCTGTCTTTCTACAAAGAAAAAGCAGAGAAGTGTAGATAATGAAGAATTACAGAACTTATGGTTTTTGTCTGGAC
AAACGCCAGACACCACCTATGATTGGAGAGGAAGCGATGATCAATTACGAAAACCTTTTGAAGGTGGTGAAGGCTGGAGCA
AAGTGAAGCAATTTTTCACAGCAAAAGTCTTTGCTAAACTCCTTCATACAGATTTCATATGGAAGAATTTCCATCATGCAGTTC
TTTAATTATGTCATGAGAAAAGTTTGGCTTCATCAAAACAAGATAGGACTCAGTTTATATGATGTCGCTGGGCAGGGGTACCTT
CGGGAATCTGATTTAGAAAACCTACATATTGGAACCTATCCCTACGTTGCCACAATTAGATGGTCTGGAATAATCTTTCTACTCC
TTTTATGTTTTGTACAGCAGTTAGGAAGTCTTCTTCTTTTATAGATCCTTTAAGAACAGGAAAGATAAAAAATCAAGATATTTTA
10 GCATGCAGCTTCCATAGATGATTTATTTGGAGCTAAGGAGTAGGAACTGTCCAAGGAGAGTCAAGAAACAAATTTGGTTTTCTGCT
CCTTCTGCCCTAAGAGTTTTATGGCCAGTACTTGAATCTTGATAAAGATCACAATGGCATGCTCAGTAAAGAAGAACTCTCACGC
CATGGAACAGCTACCATGACCAATGTCTTCTTAGACCGTGTTTTCCAGGAGTGTCTCACTTATGATGGAGAAATGGACTATAAG
ACCTACTTGGACTTTGTCTTGCATTAGAAAACAGAAAGGAACCTGCAGCTCTACAATATATTTTCAAACCTGCTGATATTTGAG
AACAAAGGATACCTGAATGTCTTTTCACTTAATTATTTCTTTAGGGCCATACAGGAACCTAATGAAAATCCATGGACAAGATCCT
15 GTTTCATTTCAAGATGTCAAGGATGAAATCTTTGACATGGTAAACCAAAGGATCCTTTGAAAATCTCTCTTCAGGATTTAATC
AACAGTAATCAAGGAGACACAGTAACCACTTCTAATCGATTTGAATGGCTTCTGGACTTACGAGAACAGAGAGGCTCTTGT
GCAATGACAGTGAAGAACTCTGCAGACCTTGATGATACATGATCTCTGAAAGACTAGACTGTCTTATATTTAGAGATACTTGAA
TGCTGCATGTAAAGCCTTTAAAGCAAAATCCTCAGAAATGGTCTAAATAAAACACTTGATATGCCTAGAGAACACAAAAA
AAAAA
20 47
MPKAKGKTRRQKFGYSVNRKRLNRNARRKAARGIECSHIRHAWDHAKSVRQNLAEMLAVDPNRAVPLRKRKVKAMEVDIEERP
KELVRKPYVLNDLEAEASLPEKKGNLTLRDLDIVRYMVENHGEDYKAMARDEKNYQDTPKQIRSKINVKRYFPAEWQDFLD
SLQKRKMEVE
48
25 GTTACACGAGGTCTGAGAGACAGAGGCAGCGTGTGTTGAGCTGCTGGTGGTGCAGCGGATGCCCAAGGCCAAGGGCAAAA
CCCGGAGGCAGAAGTTTGGTTACAGTGTCAACCGAAAGCGTCTGAACCGGAATGCTCGACGGAAGGCAGCGCGGGATCGAAT
GCTCCACATCCGACATGCCCTGGGACCACGCTAAATCGGTACGGCAGAACCTGGCCGAGATGGGGTTGGCTGTGGACCCCAACA
GGGCGGTGCCCTCCGTAAGAGAAAGGTGAAGGCCATGGAGGTGGACATAGAGGAGAGGCCTAAAGAGCTTGTACGGAAGCCCT
ATGTGCTGAATGACCTGGAGGCAGAAGCCAGCCTTCCAGAAAAGAAAGGAAATACTCTGTCTCGGACCTCATTTGACTATGTAC
30 GCTACATGGTAGAGAACCACGGGAGGACTATAAGGCCATGGCCCGTATGAGAAGAATTACTATCAAGATACCCCAAAACAGA
TTCCGAGTAAGATCAACGCTCTATAAACGCTTTTACCAGCAGAGTGGCAAGACTTCTCGATTCTTTGCAGAAGAGGAAGATGG
AGGTGGAGTGAAGTTTACATCACAGCTGCCCCAGGCTGAGGCGTCCCCGGACAGTGAAGCTGGAGCCAGGGTGAAGGCA
AGGAGTGTCTGTGGCTCCAGAGGAGCTGGCCAGGTCCCATGGAATCAGAAGGTTACACACACAGTGCACACTCCCGCTCT
GGGGAAGAACTGTCTCAGAGGCTCCAATTTATATTCATCTGGGGGTTACGGAAAGCCAGAACCTGTGTTTTTCAGGGTGG
35 GTGATGTAAATATAGTGTGTACATAATAAGCAAAATATATTTTAAAAA
49
MMPEFQKSSVRIKNPTRVEEIIICGLIKGGAALKQIITDFDMLSRFSYKGRCPCHNIIIDNCKLVTDECRKLLQLKEKYAI
EVDPLVTVEKYPYVMEWYTKSHGLLVQALPKAKLKEIVAESDVMLKEGYENFFDKLQHSIPVFIIFSAGIGDVLEEVIRQAG
VYHPNVKVSFNFDFDETGVKGFKEGELIHVFNKHDGALRNTEYFNQLKDNSNIIILGDSQGLRMDGVANVEHILKIGYLND
40 RVDELLEKYMSYDIVLVQDESLEVANSILQKIL
50
ATGATGCCAGAATTCAGAAAAGTTCAGTTCGAATCAAGAACCCTACAAGAGTAGAAGAAATATCTGTGGTCTTATCAAAGGA
GGAGCTGCCAACTTCAGATAATAACGGACTTTGATATGACACTCAGTAGATTTTCATATAAAGGAAAAAGATGCCCAACATGT
CATAATATCATTTGACAACGTGAAGCTGGTTACAGATGAATGTAGAAAAAGTTATTTGCAACTAAAGGAAAAATACTACGCTATT
45 GAAGTTGATCTGTCTTACTGTAGAAGAGAAGTACCTTTATATGGTGAATGGTATACTAAATCATATGGTTTGTCTGTTCAG
CAAGCTTTACCAAAAGCTAACTTAAAGAAATTTGTGGCAGAATCTGACGTTATGCTCAAAGAAGGATATGAGAATTTCTTTGAT
AAGCTCCAACAACATAGCATCCCGTGTTCATATTTTCCGCTGGAATCGCGGATGTACTAGAGGAAGTTATTCGTCAAGCTGGT
GTTTATCATCCCAATGTCAAAGTTGTGTCCAATTTTATGGATTTTGATGAAACTGGGGTGTCAAAGGATTTAAAGGAGAACTA

ATTCATGTATTTAACAAACATGATGGTGCCTTGAGGAATACAGAATATTTCAATCAACTAAAAGACAATAGTAACATAATTCCTT
CTGGGAGACTCCCAAGGAGACTTAAGAAATGGCAGATGGAGTGGCCAATGTTGAGCACATTCTGAAAAATGGATATCTAAATGAT
AGAGTGGATGAGCTTTTAGAAAAGTACATGGACTCTTATGATATGTGTTTAGTACAAGATGAATCATTAGAAGTAGCCAACCTCT
ATTTTACAGAAGATTCTATAA

5 51

MHTLVFLSTRQVLQCQPAACQALPLLPRELFPLLFKVAFMDKKTVVLRVLVHTWFFPLLSFQQLQECAHCSRALLQERFSTES
MQAVILGLTARLHTSEPGASTQPLCRKHALRVLDMTGLLDDGVEQDPGTMSMWDCTAAVARTCIAQQQGAEPGPAPIPVEVR
VDLRVNRASYAFLREALRSSVGSPLRLCCRDRLAEDLPMRNTVALLQLLDAGCLRRVDLRFNNLGLRGLSVIIPHVARFOHLAS
LRLHYVHGDSRQPSVDGEDNFRYFLAQMGFRFTCLRELSMSSLLSGRLDQLLSTLQSPLESLELAFCALLPEDLRLFLARSPHAA
10 HLKKLDLSGNDLSGSQALAPFQGLLQASAAATLLHLELTECQLADTQLLATLPILTQCASLRYLGLYGNPLSMAGKELLRDSVAQ
AELRTVVHFPFVDCYEGLPWPPASVLLLEASINEEKFARVEAELHQLLLASGRAHVLTWTTDIYGRLAADYFSL

52

CGGGGCTGGAGGCGGTGGCTGCGGTGCGGGACCGGCACTATGCTGGGCCTTCCTACCACTTGTGTGTGGCTTGGTAGTGGCCT
AGGGTCTCTCTCCCTGCTGAAGTCCCTCTCTGCGAGGTGGCCCTTGCCCGGCCAGCACCATGCACAGCTTGTGTCTTGA
15 GCACACGGCAGGTGCTGCACTGCCAGCCAGCTGCCTGCCAGGCCCTGCCCTGCTGCCACGCGAAGCTTTCCTCCCTGCTGTTCA
AGGTGGCCCTCATGGACAAGAAGACAGTGGTACTGCGGAGTTGGTACACACGTTGGCCCTTCCCGCTGCTCAGTTTCCAGCAGC
TGCTACAGGAGTGTGCCCACTGCAGCCGTGCCCTCTGCGAGGAGCGGCTAGCACTGAGAGCATGCAGGCTGTTATCTTGGGGC
TGACTGCCCGGCTCCACACCTCAGAGCCTGGGGCCAGCACACAGCCCTCTGCGAGGAAGCATGCGCTGCGGGTGTGGACATGA
CGGGCCTCTTGGATGATGTTGTGGAACAGGATCCTGGCACCATGAGCATGTGGGACTGTACTGCTGCCGTAGCTCGCACATGCA
20 TTGCCCAGCAGCAGGTGGGGCCGAGAGCCTGGGCCAGCCCCATCCCGTGGAGGTGCGCGTGGACCTGCGGGTGAACCGGG
CCTCCTATGCGTTCTGCGGGAGGCACTCCGAAGCAGCGTGGGCAGCCCGCTGCGGCTCTGCTGCCGGGACCTGCGAGCTGAGG
ACCTGCCCATGCGCAACACTGTGGCCCTGCTGCACTTCTGGATGCGAGGCTGCCGTGCGCGCGTGGACCTGCGCTTCAACAATC
TGGGCCCTGCGCGGCTGTCTGTGATCATCCACAGTGGCCCGCTTCCAGCACCTGGCCAGCCTGCGGCTCCACTATGTGCATG
GGGATTCAGGCAGCCCTCCGTGGATGGCGAGGACAACCTCCGCTACTTCTTGGCCAGATGGGCCGCTTACCTGTCTGCGTG
25 AGCTCAGCATGGGCTCTCTCTCTTTCAGGGAGGCTGGACAGCTGCTCAGCACCTTGCAGAGCCCTGGAGAGCCTGGAGT
TGGCCTTCTGTGCTGCTGCTGCTGAGGACCTACGCTTCTGCGCAGGACCCACATGCTGCCACCTCAAGAAGTTGGACCTGA
GTGCTAACGACCTGTCTGGCAGCCAGCTGGCACCTTCCAGGGTCTGTGTCAGGCATCAGCAGCCACACTGTTGCATCTGGAGC
TGACTGAGTGTGAGCTCGCAGACACCCAGCTGTTGGCCACACTACCCATCTGACTCAGTGCAGGAGTCTCCGTAACCTTGGCC
TCTATGGCAACCCACTGTCCATGGCGGGCTCAAGGAGTGTGCGGGACTCAGTGGCACAGGCTGAGTGCCTACTGTGGTGC
30 ACCCTTCCCTGTGACTGCTATGAGGGCTTGCCTGGCCGCGGCTGCTCTGTCTGCTGGAGGCTCCATCAATGAGGAGA
AGTTTGGCCGCTAGAACTGAGTTGCACAGCTGCTTCTAGCCTCAGGCGTGGCCATGTGCTCTGGACACGGACATCTACG
GGCGACTGGCTGCGGACTACTTACGCTATGATGAAGTAGCTCTGGGTGAGACACAGGCCGCTTGCAGTCTCTTAGTAGGC
AGGGCCTTTGCTGGGACCCCTGGTGGAGGCTTCAAAAAGCACTGGTTACTGGTTCTTGTGGGTCTACCTTGTCTTGGGC
ACACCTCAAGCCTCCCTGCTTCTGCGAGTGGCCACGCGGTTTCCCTGCACCTGCTCCATAATTGGTGATCATCTGTGGGC
35 CCGGGGCTGGATGTCAGGCTCCATTGCGCTGCTCAGTTTGGGTGCATTTGGCTGCCGTCTGGGGTCTTGGTCTTTGTGCAA
ATGCTTTGGGATTCAGTTGTGAGCTGAGAGAGATGATGCCCTCTTGGGCCCTTCTCTCCCTTACTGAGAGCTCAGTGCTT
CTGGGGTTGAAGTTGGACAGAGGCTGCTTCAGGGAAGCTGGGAGTCCCAGGCACTCAGCCTCTTACGTGTTCCCTACCCCTG
CCACCCAGCCACCCACTGGGCTGGGCTCGGGCTGGAGGGGTGATCAAGGTACACATGTGCTGGAAGTTGAATTTGTGGCTG
TTTCTTTTCTGTCCACGTTGGTCACCTTATCTTATCTCTGCTGTACCCCCAACATGGCCACCGGGCAACAATGCCATCC
40 AGCCTGTGCGCCCGCCCTTTCGCGGGCAGCCCGTGGCACTGCCCGGCACTGCTTGTCTTCCCACTTTCGGAGGCCAAGA
TCTTACTGTGGCCGGCCAGGGCAGCGAGGACCCCCCATGCGAGCTGGAGGTGGGGTGATGCTTTTCGGAAGAGCTTC
AAGGGAGGTGTTGGGGCTTCCCGGCCACCTTCCATTGCTACCCAGGATTCGGAGTGAACGTTCCCGGCTCGCGCCCCACA
CAGCGCTCAGCGCACACTGCGCGGCTTCACTTTACTGACGGAGCATGCGCGAGGCCGACCGGCCAATCTCCGGCGCCCCACG
TCATCCGCGCGCCCGGGCCCTAGCAGTGGATCTTAGGCGACCGCGGGGACGCGGAGTCCCGGGCCCGGCTGTTCTTCC
45 GGGCCGAGTCAAGCGGGCCTCCGCCGACCTCGCGGAAGAGCGGCTTGGAGCGGTTGATGACGAACATCTGTTGGCGGCG
TCGTCGCGGAGCTCCTCTAGCTGTGCGAAGCTACAGGGCGCTCCAAGTAGTGTGACGAACAGCGCTCCCTCCCCGGAGGC
CCCCCGGCTTTTTCGCTGCGGCGCGGCGACAGATCATGGCGACAGGAGCAGCGCGTGAGCGCCAGCAGCGGATGGCC
GCCGCAATGGCGTCTGTGTGGCCACGCCAGGGCGGGAAGGCCATGTGCGCCGCTCGGGCCGGGGTCTGCTGCGGCGGGG
CGGGCGGCGGAGCGGCGGTTGCGCGGCTGCTGCGGCTGCTGCGGGACGCGTTGACCAGGAGCGGAAGGGCACGCGGGCA
50 GCGCGCGCGGCTTGGAGGCTCGCACTCGTACTTACCGCGTGGCCAGCGTGATGTTGCTGAGGAAGAGCATGCCGTGCCC

- GTGTGGATGCCGAGTGTCCGCCAGGCCAGCAACCCGCCTTCTAGCTGGGCCCTGGGCTCGCGGCCGCCCTCGCGAGGCTGG
GGCACCCTTTCTCCAGGTACCAATGGCTGCGGGTAGCCGGAGGCTTGGCAGGCAACCCGAGGTCTCACCAGGTTGGCTGTG
AGCTCCAGCGGCTGCACGTGGACAGAGGGCGGAATGCAGATGAGGCTGCTGTGGGATACGTCCAGGAGACTCTGGAGCGCCAGG
CGCGGGGGCTCTGCACACATGATCTTCTGTCCCTGGAGGTGAGCAGCCGCTGGCCGCCCTCCTTTGATCCAGGCCCCCAGCCAG
5 TGCAGGGCGCAGTCACAGCGCCATGGGTTCCTGTGTGGGAGAGCAGCGTTAGGCAGGTGGCTTGAGGTGTGCTAAACAGCCT
GTGCAGTTGGGGTTTTCAGGCCAGGACAGAGGCCCTCTTTCCACCTCCACAGCGTTTTCACACGGAGTCCAAGGCCCTGCCA
CCCCCTTCTTGACCCCAAGCTCTTGGGGCGGAGGCCCTTACCCCTCGCCCCCTCCCTTTAGCTCTGTGATGCCTGCCTGT
TACAGATCACTTCTCCGTCGGTCTCTGAGAAAGCACCTGCTCTTAAAGTCTTCTGCAACAAGTGCCACTGTTTTAGGAACCT
GGGCGTCCACATAGACATCTCACCAGCACTGAAACCTCACAGTCTCTCAGCTTTGCTTTGGATGCCCTCTCTTGGGAATGT
10 CCCCAGTCTGGTCAGCTGTCTCTCTCTTTGCAATTTGTCTGCCTCCCTCAGCCTAAAAGTGTGCAGAACCTCAATCTGT
TTAAGTCACCCTGTGGAGTTCCTGTCTTCTGTTTCCCGAGGAGGCTGCTGAGCTGTATCCCCAGCACACCCACTCCCGCA
GCCCCTCCAGCTGTGGAGTTCCTGTCTTCTGTTTCCCGAGGAGGCTGCTGAGCTGTATCCCCAGCACACCCACTCCCGCA
CCTGTCTCCCTACTGCCAACACCGTGGCCCGAGACAGAGACGCTTTCAGAGGAAGAGGTACCTGTGAGGCGCAGGACTTGCAGACT
GGCCAGGGGCTGCAGGGCTCTCGCTGATGGTGCAGGCTGGTTCTGTGAGGTCCAGCAGTGTAGGAGGAGCAGCCCGC
15 TAGAGCCTGGTCTCCAGCAGCTCAATGCTGTTTTCTTGCAGGTGAAGCTCCTGCAGTCGCTGAAGTAAGGACAGCAGATCGTG
AGGAAAAAGGGCGCCGAGGTGGGGGCATGTCTCTCTTCTTACCAAGCTAGACTGGGTGGCTTTTCTAATATTCCAGCCCTA
CAGGGCGAGGGCCATAATGGAGTATCCCGCCCTTTAGACCCAGGCGCTACCGGCAGGTGCAAGAAGGTGAAATCCAGCAG
CCGCGCCAGCTGGTTGCCCGCCAGGTAGAGCAGCGCAGCTGGGCCAGGCTACGAAGGCGCCGCTGCGCAAGCCGCGCAGCCG
GTTGCTAGTGAGCGCCAGCTCCAGCAGGCGCGGCTGCGCGCGGAAGGCGCGGCCCTCCAGGCGCGCAGGCTGTTGTTGTGCAG
20 GTAGAGCCGCGCAGAGCGCGGAGTGGCGCAGGCTCCCGGCTTAGGCGGGCGATGTTGTTGTCTGCAGGAACAGTGTCTG
CAGGCGGGGAAAGAGGAGGCGCTTACCCGCTGCGGGGTCTTCTCTCCTTGGGGGACCCCGCTCTCCCGAGCTCCACAC
GGTCCCACCTGCGTCCCTGGCGGATTCACAGCGGACGACGCGCAACCGCAGGGCGCCACACTCCACCTGGCGCTGTAGCA
GCGGCAGGCTGCTGGGCAGCGCGCGCGGAGCGGCAGTAGTAGCAGCAGCAGGCGCAGAGTTGCGGGGCCCTCAGGGCCAT
CTCCCGAGGCGCGGTTCTTACCGGCCCTTCCCGGTTACGCGCAGACGCGTGCCTCCTGAAACACAGGTTGGCAGGCCAGT
25 CTCGGCAGTCGAGAGCCAGCCAATAGATGGAATGGAGGCCCTGCACCTGCGTCTAACTTTTGACACTATAAATAGGTTCAAGAAA
CTAATAAACGTTCTGGTTTCTCCTTTGAC
- 53
- MLQTSNYSVLVLSLQFLLSYDLFVNSFSELLQKTPVIQLVLFIIQDIAVLFNIIIFLMFFNTFVFAQGLVNLFFHKFKGTIIL
TAVYPALSI SLHVWMNLRWKNSNSFIWTDGLQMLFVFQRLAAVLYCYFYKRTAVRLGDPHFYQDSLWLRKEFMQVRR
- 30 54
- ATGCTCCAGACCAGTAACACAGCCTGGTGTCTCTCTGCAGTTCTGTGCTGTCTATGACCTCTTTGTCAATTCCTTCTCA
GAATGCTCCAAAAGACTCTGTATCCAGCTTGTGCTCTTCATCATCCAGGATATTGAGTCTCTTCAACATCATCATATT
TTCTCATGTCTTCAACACCTTCGTCTTCCAGGCTGGCTGGTCAACCTCCTATTCCATAAGTTCAAAGGACCATCATCTG
ACAGCTGTGTACTTTGCCCTCAGCATCTCCCTTCATGTCTGGGTATGAACCTACGCTGGAAAACTCCAACAGCTTCATATGG
35 ACAGATGGACTTCAAATGCTGTTTGTATTCCAGAGACTAGCAGCAGTGTGTACTGCTACTTCTATAAACGGACAGCCGTAAGA
CTAGCGCATCTCACTTCTACCAGGACTCTTTGTGGCTGCGCAAGGAGTTCATGCAAGTTCGAAGGTGA
- 55
- MGGFQRGKYGTMAEGRSEDNLSATPPALRIILVGKTGCGKSATGNSILGQPVFESKLRAQSVTRTCQVKGTWNGRKVLVVDTP
SIFESQADTQELYKNIGDCYLLSAPGPHVLLLVQLGRFTAQDTVAIRKVEVFGTGAMRHVVILFTHKEDLGGQALDDYVANT
40 DNC SLKDLVRECERRYCAFNNWGSVEEQRQQAE LLAVIERLGREREGSFHSNDLFLDAQLLQRTGAGACQEDYRQYQAKVEWQ
VEKHKQELRENE SNWAYKALLRVKHLMLLHYEIFVFLLLCSILFFIIFLFIFHYI
- 56
- GTTCCTCAATCAGTTTCCAGCCAACACCAGGCTGCTCTAGTCCGAGAGGTGTGGGGGACACACTCCATAATCTCTACTTTTCTT
TTTGTGCAGCTGAGTCATGGAGCTTTCAGCCCCAGCACATGGCTCTCTTAACTGCGTCTGCTCAACCTCCCTCAGCCCTGTG
45 AACAGCATCCCCGACACAGACGAGCAGGACTCTCTGCTGCCACTTCACCTTCTGAGAGAGGACCAGCGCCAGAGCC
TCAGTGAAGTCCACCTTGGAGGACAGGGCACAACACCGTTTCTGAAGAGAATGGGAGGATTCCAGAGGGGCAAAATATGGAAC
ATGGCTGAAGGTAGATCAGAAGATAACTTGTCTGCAACACCACCGCATGAGGATTATCTAGTGGGCAAAACAGGCTGCGGG
AAAAGTGCCACAGGGAACAGCATCTTGGCCAGCCGCTGTTGAGTCCAAGCTGAGGGCCAGTCACTGACCAGGACGTGCCAG

GTGAAAACAGGAACATGGAAACGGGAGGAAAGTCTGGTGGTTGACACGCCCCCATCTTTGAGTCACAGGCCGATACCCAAGAG
CTGTACAAGAACATCGGGGACTGCTACCTGCTCTCTGCCCCGGGGCCCCACGTCTGCTTCTGGTGATCCAGCTGGGGCGTTTC
ACTGCTCAGGACACAGTGGCCATCAGGAAGGTGAAAGAGGTCTTTGGGACAGGGGCCATGAGACATGTGGTCATCTCTTCACC
CACAAAGAGGACTTAGGGGGCCAGGCCCTGGATGACTATGTAGCAAACACGGACAACATGCAGCCTGGAAGACCTGGTGGGGAG
5 TGTGAGAGAAGGTACTGTGCCTTCAACAACCTGGGGCTCTGTGGAGGAGCAGAGGCAGCAGCAGGAGCTCTGGCTGTGATT
GAGAGGCTGGGGAGGGAGCGAGAGGGCTCTTCCACAGCAATGACCTCTTCTTGGATGCCCAGCTGCTCCAAAGAACTGGAGCT
GGGGCTGCCAGGAAGACTACAGGCAGTACCAGGCCAAAGTGGAAATGGCAGGTGGAGAAGCACAAAGCAAGAGCTGAGGGAGAAC
GAGAGTAACCTGGGCATACAAGCGCTCTCAGAGTCAAACACTTGATGCTTTTGATTTATGAGATTTTTGTTTTTCTATTGTTG
TGCAGCATACTTTTTTTCATTATTTTTCTGTTTCATCTTTTTCATTACATTTAAATCTCTGGACCTGGAGCACTTCTAATGTATCA
10 CCCCATGGAGTCATGTTCTAATAATACCAATTCAGACTCAGATCCTCGTGGTCTATGGAGCATGCTGCTTGTGCTGTGTGCA
GCTCCCATTTCCCTTCTTCCCTGATAGACTTGGAGCTGTGTGCCCTCCACTCCAAGGCTGCCTGCCTGCTAAGCACTATTCCA
CTCTGTCTGCCAACAACTGCTTCAAGAAATGGGCCTGAGATCCCATGCAGGTCCCTGAGAAGTGAGTAAAGTCCGCAGAGGTGG
GGATGGAAGATCTCTCTTAGATAGAACCTGTCTTCTCCCTGGCATTTGTGGGTCTGGGCGTGACACTGGGACTCTCAGCAGC
TTTGTGCTGCCAACCTGAGATTGAAGGCAGTGCCTCAGAGCAGCACAGAGATTGGGGCCCCCTGAGCCCTGAGCCACCAGCCC
15 TGCAGCCTGCCCTATCTCCGATTTCCAGTTGTATTAGCCAATAGATTTCCTACTTATTTAAGCTATTTGAGCTCCGGGTCTCT
TCTACCTGCATTCTAAACATTCAAAGTAATAAAAAATTTCTCCACATT

57

ELGIPALLDPNDMVSMSVPDCLSIMTYVSQYYNHFCSPGQAGVSPPRKGLAPCSPPSVAPTPESEDAQGEELSSGSLSEQGT
GQTPSSTCAACQQHVLVQRYLADGRLYHRHCFRCRRCSSTLLPGAYENGPEEGTFVCAEHCARLPGTRSGTRPGPFSQPKQQ
20 HQQQLAEDAKDVPVGGPSSAPAGABADGPKASPEARPQIPTKPRVPGKLQELASFPAGRPTAPRKASESTTPAPPTPRPRSS
LQQENLVEQAGSSSLVNGRLHELVPVPRGTGTPKPSGTPAPRKDPPIWTLVQAEPPKKPAPLPPSSSPGPPSQDSRQVENGTE
EVAQPSPTASLESKPYNPFEEEEEDKEEEAPAAPSLATSPALGHPESTPKSLHPWYGITPTSSPKTKKRAPRAPSPASPLAHA
SRLSHSEPPSATSPALSVELSSESASQTAGAELEPPAVPKSSSEPAVHAPGTPGNPVSLSTNSSLASSGELVEPRVEQMPQ
ASPLAPRTRGSSGQPAKPCSGATPTPLLLVGRSPVSPSGSSSPQLQVKSSCKENPFNRKPSAASPATKKATKSKPVRPP
25 APGHGFPLIKRKVQADQYIPEEDINGEMDTIERRLDALHRRGLVLEEKLRLGGLNEGREDMDLVDFKLIHEKHLVLRRESELIY
VFKQQNLEQRQADVEYELRCLLNKPEKDWTEEDRAREKVLMLQELVTLIEQRNAIINCLDEDRQREBEEDKMLEAMIKKKEFQRE
AEPEGKKKGFKFTMKMLKLLGNKRDASKSPRDKS

58

GGAGCTGGGGATCCCCGCTCTCTGGACCCCAATGACATGGTCTCCATGAGCGTCCCTGACTGCCTCAGCATCATGACCTATGT
30 GTCCAGTATTACAACCACTTCTGCAGTCTTGCCAAGCTGGTGTCTCGCCACCCAGAAAGGGCCTTGACCCCTGTTCCCGGCC
GTCTGTAGCACCCACTCCAGTGAATCAGAAGATGTGGCTCAGGGCGAGGAGCTCTCTCAGGCAGCCTGTGAGCAGGGCAC
CGGCCAGACCCCCAGCAGCAGCTGCGCAGCCTGCCAGCAGCATGTGCACTTGGTGCAGCGCTACCTGGCTGACGGCAGGCTGTA
CCATCGCCACTGCTTCCGGTGTGCGCGGTGCTCCAGCACCTGCTCCCCTGGGGCTTATGAGAATGGGCCTGAGGAGGGACCTT
TGTGTGTGAGAACTGTGCCAGGCTGGGCCCGGGACACGGTCCGGGACCAGGCTGGGCCCTTCTCAGCCAAAGCAGCA
35 GCACCAGCAGCAACTCGCAGAAGATGCCAAGGATGTTCCAGGAGCGGGCCCCAGCTCCAGTGTCTCTGAGGGGCTGAGCCGA
TGGACCCAAAGGCCAGCCCTGAGGCCCCGGCCGAGATCCCTACCAAGCCCCGGGTTCCTGGCAAACATACAGGAGCTGGCCAGCCC
CCCTGCGGGCCGCCCCACCCCTGCCCCAGGAAGGCCTCTGAGAGCACCACCCAGCACCACCCAGCCCCGGCCCCGCTCCAG
TCTGCAGCAGGAGAACCTGGTGGAGCAGGCTGGCAGCAGCAGCTGGTGAACGGGAGACTGCACGAATGCCTGTCCCCAAGCC
GAGGGGGACACCGAAGCCGTCCGAGGGGACACAGCCCCAGGAAGGACCCCCATGGATCACGCTGGTGCAGGCAGGAACCAAA
40 GAAGAAGCCAGCCCCACTTCCCCAAGCAGCAGCCCCGGGGCCACCAAGCCAGGACAGCAGGCAGGTGGAGAATGGAGGCACCGA
GGAGGTGGCCAGCCGAGCCCAACGGCCAGCCTGGAGTCCAACCCCTATAACCCCTTTGAGGAGGAGGAGGAGCAAGGAGGA
AGAGGCTCCAGCTGCACCCAGCCTGGCCACAGCCCTGCCCTGGGCCACCCGGAGTCCACACCAAGTCCCTGCACCCCTGGTA
CGGCATCACCCCTACAGCAGCCCCAAGACAAAGAAGCGCCCTGCCCGCGCGCACCCAGCGCGTCCCCACTGGCTCTCCACGC
CTCCCGCTCTCGCACTCGGAGCCGCCCTCGGCCACACCATCGCCAGCGCTCAGCGTGGAGAGCCTGTGCTGTGAGAGCGCCAG
45 CCAGACTGCAGGTGCAGAGCTTCTGGAGCCGCCAGCTGTGCCCAAGAGCTCCTCAGAGCCTGTGTCCATGCCCTGGTACCCC
TGGAAACCTGTGACCTCTCTACCAACTCCTCCCTGGCTCTCTGGGGAACAGTGGAGCCTAGAGTGAACAAATGCCTCA
AGCCAGCCCTGGCTTGGCCCCAGGACCAGGGGAGCTCAGGTCCCCAGCCAGCCAAGCCCTGAGTGGGCCACCCCAAGCC
TCTCTTGTGGTTGGAGACAGGAGCCCGTGCCTTCCCCTGAAGCTCGTCCCCACAGCTGCAGGTAAAGTCTCTGCAAGGA
GAATCCTTTTAACCGGAAGCCATCAGCTGCAGCTCCCCAGCCACAAAGAAGGCCACCAAGGGATCCAAGCCAGTGGGCCACC
50 TGCCCCTGACACGGCTTTCCACTCATCAAACGCAAGGTCCAGGCTGACCAGTACATCCCTGAGGAGGACATCCATGGAGAGAT

GGATACCAATTGAGCGCCGGCTGGATGCCCTGGAGCACCGTGGGGTGCTGCTGGAGGAGAAGCTGCGTGGCGGCCTGAATGAGGG
CCGTGAGGATGACATGCTGGTGGACTGGTTCAAGCTCATCCACGAGAAGCACCTACTGGTGGCGGAGAGTCCGAGCTCATCTA
TGCTTTCAAGCAGCAGAACCTGGAGCAGCGCCAGGCTGATGTCGAGTATGAGCTCCGGTGCCTCCTCAATAAGCCAGAAAAGGA
CTGGACCGAGGAGGACCGGGCCCGGAGAAAGGTGCTGATGCAGGAGCTTGTGACCCTCATTGAGCAGCGCAACGCTATCATCAA
5 CTGCCTGGATGAGGACCGGCAGAGGGAGGAAGAGGAAGACAAGATGTTGGAAGCCATGATCAAGAAGAAAGAGTTCCAGAGGGA
GGCTGAACCTGAGGGCAAGAAGAAGGGGAAGTTCAAGACCATGAAGATGTTGAACTGCTAGGAAACAAACGTGATGCCAAGAG
CAAGTCCCCCAGAGACAAGAGCTAACAGCACGAGAAGCCAGTTGGGGACTGCCCCCTCCTGGAGCAGCTCCTGGGCTGTGCTCT
GTTTGAAGGGGGCGCCCTGCTCCCTCAGATCAGTCAGGAGGAAGATGACTAAGGGGAGGGATCCTCTGGGTGATGGCCCTCTTC
CTCCTCAGGACCTCTGACTGCTCTGGGCCAAAGAATCTCTTGTCTCTCCGAGCCCCAGGCAGCGTGATTACGCCCTGCC
10 CAACCTGATTCTGATGACTGCGGATGCTGTGACGGACCCAGGGGCAAAAGGGTCCAGGGTCCAGGGAGGGGCGCCCTGCTGA
GCATTCCGCCCCCTCACCTGCCAGCCCCCTGCCATGAGCTCTGGGCTGGGTCTCCGCCCTCCAGGGTTCTGCTCTTCCAGGCAG
GCCAGCAAGTGGCGCTGGGCCACACTGGCTTCTTCTGCCCATCCCTGGCTCTGAGTCTCTGTCTTCTGTCTGTGCAGGCG
CCCTTGGATCTCAGTTTCCCTCACTCAGGAACCTGTCTTCTGAAGTCTTCAGTTAAGTTTGAAGTTTATGACTGAGTGGCTGT
CTGTGACAGCTGAATGGGCTGACGGGCAATCCATCCCTCTCTCCCTCACAGTTCCAGGAGCGGCTTCCCTCGTCTCCCTTA
15 CTCCACAGGAGCCCTCCCTTGGCAGGACAGGGCTGCGACGGCCATGCTGGGGCAGGTGAGTGTCTGTAGTGTCTCCAGTG
CTGTCCCCAGGCTGCAGTTCTGGTCCCTGGTTGTGACGTTAGGAAGGGTGCAGTTGAAGCAGGTGCTCATCTCGGTTCTTAAACG
TTTATAGTCTGACCCCTCACTTAGGCTTTCTCTGCCACCCCGTCCAGGAAGAGGCTCGCTCCCGCCATGGTCATCACTGG
TCTGTCTGTCTGTGTCTGTCTTCTTCCCTGACTCCCTCCACCGAAGGCTGATGGCTACTCACCCCTCTGGGATGGCTATGG
GAGAGGAGGAGTGTGAGGACCGCCACCTTTCTGACGAAATGTGCCAGCAGCTCTTGGTCAAAGCAGTGTGTCTATAAGCT
20 ATCTCTGGGATGCCTCTAGGCCCCCTTCCCTCTACACACCTCTGGGAAAAGATTACACTGTATTAACTCTCGAGGAGTTTCCCTC
ACCAATAAACAGACAACCTCAACTGCCAGTGCCCTGCAGCTCGGGCCACAGCGGCAGCCTTGTGTGCTTCCACCTGCCTCT
GCCACACCTGGTGGCTGAACATCTCTGGTGGCCAGAGGCCATGTTGGGGCCATCTCCAAGAGGGATCTTGCCCTCACCGCC
TGCCACTGGGCAGGATCCCTTTCTCTGACGGGAGAGGTGGCTCCTCGGCCATGCAGCCCTGGCAGGCTCCTTCTAAACATGC
CTGTTGACCTGGAGCTGGCGCCACCACTCCAGGGCTTTCAGGGCCAGACAGGTAACACGCATGAACCCGAGTGACAGCTCT
25 GACGGGCTGTTTCGGTGTGACGAGACAAAGCTGGCAGGGCAGGGGTGAAGTGGAGGCAAGTCAAGTCACTGTGGCTGTGGG
GTGAATGTGGGCCCGGTGTGTCAGATCCTTTGTCATAAGAAGCTAGAAATCCAGATTTTATGTGTGTAAATTTGTAATGTC
TGAAAGCTAGCCTGAATTTTCTTTTGTGAGACAGAGTCTCGCTCTGTGCGCCAGGCTGGAGTGACGTGGCGCATGCTCA
GCTCACTGCAGGCTCCGCTCTCTGGGTTACGCCATCCTCTGCTCGGCTCCTGAGCAGCTGGGACTACAGGCGCATGCTAC
GACGCTGGCTAATTTTGTATTTTGTAGTAGAGACGGGTTTACCCTGTTAACCAGGATGGTCTCGATCTCTGACCTTGTG
30 ATCCACCCACCTTGGCTTCCCAAAGTGTGGGATTACAGGCGTGAGCCACCACGCCCCGGCCACTAGCCTGAATTTCAATCAAGG
GTTGGCTGATACTGTGTGTCAGGGTGGACTGGATTTGTCTGGGGGTCTCTGGTTGTGCTGCTCTGACCACATGATGGGG
CCTTCGAGGTCGAGGACAACCTGTTCCATTAGATTGCACCTCTGCCCTCAGGTTCTTGAGGGTGTGTGGACACAGAGGCTTTC
CATGGGATGTCCCTGAGCCGGCCCTTGATTGGGGCTCACCATTACAGGGCCGTTTATTTCTGCAAACCGAACTTGGGTCT
GTGACCTGATGGGATTATGGGACTCCCTCCAGGTGCCGAGACAAGGTTGATATTTCCAAAATATTTTGGTGATTAGTGGGAC
35 AAGCAAATGACAGAATACCGGAGAAGGCAGGGATCGTGGGTGTGACGAGCCAGAGGGAGGGGACAGATGTGCTGTGTACAGG
ACAAGGTGTGAGGTGACTCCTTCCAGCAGGGCCTCGCAGATGCACAAGCACGGAGCTGGTGGGTTTGGCCAAAGAAAGGTCAC
GCGGCACATGCAGGGATTGGAACCTCCAGGCCAGGGCTCTAGGTGCTCCACCTTTTCATGTTTCTTCTGTGGCCATGGGTA
TAGTGGAAAGACATAAAGCTAAAGCCAACCTTTAATCTGAATGCACTGCTTGCCAGGTAATGCCCTTGGTTGTGGTATCTTG
TTGAGACTTAGTTTTCACAGAGGGATAATGAACCGTTGCAGAGGTTTATGAGATCATTAACAGACTGGAATTCAGCACCTGCC
40 ACTGCATCTCAGCTGGGCGACAGAGCAAGACTCAGTCTCAAAAAAAAAAAAAAATAGCTGGGCAGGGCATGGTGATGGTGCT
GTAATCCTAGCTACTTGGGAGGCTGAGGCAAGAGATTTGCTGAAACCCAGGAGGTGGAGGTTGAGTGAGCCGAGATCGTGCCA
CTGTACTCCAGCCTGGGTGACAGCGGAGACTCCGTCTCAAAAAAGCTGGGTGTGGGGAACACCTGTGGTCCCAGCTATTTCT
GAGACTGAGGCAGGAGGATTGCTTGAGCTCAGGAGTTCTGGCTGCAGTGAGCTATGATCATGCCACTGTATTACAGAATGGGTG
ACAGAATGAGAGCGACACTGTCTCAAAAAAAGGCGGAGCGGTCGTTGTGCTGTAAATCCCAACACTTTG
45 GGAGGCCAGGGTGGGCGGATCACTTGAGGCCAGGAGTTCAAGACCAGCCTGGGCCAACATGGTGAATCCCCATCTCTACTAAAAA
AATTAACGTGACATGGTGGTGGACACTTGAATCCAGCTACTCAGGAGGCTGACACATGAGAATTGCTTGAACCCGGGAGGCG
GAGGTTACAGTGAGCCGAGATAGCACCACTGCATCCAACCTGGGCACAGAGTAAGGCTCTGTCTTT

59

MASEGPREPESEGIKLSADVKEFVPRFAGLNVWLESSEACVFPSSAATYYPFVQEPVTEQKIYTEDMAFGASTFPPQYLSSE
50 ITLHPYAYSPTYLDSTQNVYSVPGSQYLYNQPSYRGFTVKHRNENTCPLQEMKALFKKRTYDEKKTYDQKFDSEADGTI
SSEIKSARGSHHLSIYAENSLKSDGYHKRTDRKSRIIAKNVSTSKPEFEFTTLDPELQGAENNMSEIQKQPKWGFVHSVSTDI

- SLLEVVVKPAAVLSKGEIVVKNNPNESVTANAATNSPSCSTRELSWTPMGYVVRQTLSTELSAAPKNVTSMINLKTIIASSADPKN
VSIPSSSEALSSDPYNKEKHIHPTQKSKASQGSLEQNEASRKNKKKKEKSTSKYEVLTVQEPPIEDAEFPNLAVASERRD
RIETPKFQSKQPPQDNFKNNVKKSQLPVQLDLGGMLTALEKKQHSQHAQSSKPVVSVGAVPVLSKECASGERGRMSQMKTP
HNPLDSSAPLMKKKGQREIPKAKKPTSLKKIILKERQERKQRLQENAVSPAFTSDDTQDGESGGDDQFPEQAELSGPEGMDLI
5 STPSVEDKSEPPGTELQDTEASHLAPNHTTFPKIHSRRFRDYCSQMLSKEVDACVTDLKELVRFQDRMYQKDPVKAKTARR
LVLGLREVLKHLKLLKLCVLIISPNCCKIQSKGGLDDTLHTIIDYACEQNIFFVFALNRKALGRSLNKAVPVSVVGIFSYDGAQ
DQFHKMVELTVAARQAYKTMLENVQQLVGEPRPQAPPSLLTQGPSCPAEDGPPALKEKEEPHYIEIWKHLEAYSGCTLELEE
SLEASTSQMMNLNL
- 60
- 10 .CCAAGCCGACGGCCCGCTGCTGGCCCTCCGTGACGCGCCCTCCTCCGCGCCTCGCGGCATGGCGTCGGAGGGCCGCGGGAGCCC
GAAAGCGAGGCATCAAGTTATCAGCAGATGTCAAACCATTTGTCCCAGATTTGCCGGGCTCAATGTGGCATGGTTAGAGTCC
TCAGAAGCATGTGTCTTCCCCAGCTTCGACGCCACATATCTCCGTTTGTTCAGGAACCACAGTGACAGAGCAGAAAAATATAT
ACTGAAGACATGGCCCTTGGAGCTTCAACTTTTCCACCTCAGTATTTATCTTCTGAGATAACTCTTCATCCATATGCCATTTCT
CCTTATACCTTGACTCCACACAGAATGTTTACTCAGTGCCCTGGCTCCAGTATCTTTATAACCAACCCAGTTGTTACCGAGGT
15 TTTCAAACAGTGAAGCATCGAAATGAGAACACATGCCCTCTCCACAAGAAATGAAAGCTCTGTTTAAAGAAAAACCTATGAT
GAGAAAAAACGTATGATCAGCAAAAGTTTGACAGTGAAAGGGCTGATGGAACATATCATCTGAGATAAAATCAGCTAGAGGT
TCACATCATTTGTCCATTTACGCTGAGAATAGTTTGAAATCAGATGGTTACCATAAGCGAACAGACAGGAAATCCAGAATCATT
GCAAAAAATGATCTACCTCCAAACCTGAGTTTGAATTTACCACACTGGACTTTCTGAACTGCAAGGTGCAGAGAACATATG
TCAGAGATACAGAAGCAACCAAGTGGGGACCTGTCCACTCTGTCTCTACCGACATTTCTCTTCTAAGAGAAGTAGTAAACCA
20 GCTGCAGTGTATCAAAGGGTGAAATAGTGGTGAATAAACCCTAATGAATCTGTAACGCTAATGCCGTACCAATTTCTCCT
TCATGTACAAGAGAGTTATCTTGGACACCAATGGGTTATGTTGTTTCGACAGACATTTATCTACAGAACTGTCTCAGCAGCCCCATAA
AATGTTACTTCTATGATAAACTTAAAGACCATTTGCTTCATCAGCAGATCTAAAAATGTTAGTATACCATCTTCTGAAGCTTTA
TCTTCGGATCTTCTTACAACAAAGAAAAACACATTATTCATCTTACCCTAAAGCTTAAAGCATCACAAGGTAGTGACCTTGAA
CAAAATGAAGCCTCAAGAAAGAAATAGAAAAAGAAAGAAATCTACATCAAAATATGAAGTCTGACAGTTCAAGAGCCTCCA
25 AGGATTGAAGATGCCGAGGAATTTCCCAACCTGGCAGTTGATCTGAAAGAAGAGACAGAAATAGAGACACCGAAATTTCAATCT
AAGCAGCAGCCACAGGATAATTTAAAAATAATGTAAAGAAGAGCCAGCTTCCAGTGCAGTTGGACTTGGGGGGCATGCTGACA
GCCCCTGGAGAAGAAGCAGCACTCTCAGCATGCAAAAGCAGTCTCCAAACAGTGGTAGTCTCAGTTGGGAGCAGTCCAGTCTCT
TCCAAAGAATGTGCATCAGGGGAGAGAGGGCCGCGCATGAGTCAATGAAGACCCCGCACAAATCCCTTGGACTCCAGCGCCCCA
CTGATGAAGAAAGGGAAGCAGAGGGAGATCCCCAAGGCCAAGAAGCAACCTCACTGAAGAAGATTATTTGAAAGAACGGCAA
30 GAGAGAAGCAGCGTCTTCAAGAAAATGCTGTGAGTCCAGCTTTTACCAGTGATGACACACAAGATGGAGAGAGTGGTGGTGAT
GACCAGTTTCCCGAGCAGGCAGAGCTGTGAGGGCCAGAGGGATGGACGAAGTATCTCCACTCTTCCGTTGAGGACAAGTCT
GAAGAGCCACCAGGCACAGAGCTCCAGAGGACACAGAGGCTCCACCTTGCTCCCAATCACACCACCTTCCCTAAGATCCAC
AGCCGAGATTTCAGGATTACTGCAGCCAGATGCTTAGTAAAGAAGTGGATGCTTGTGTTACCGACCTACTCAAAGAACTGGTC
CGTTTCCAAGACCGTATGTACCAGAAAGATCCAGTCAAGGCCAAGACTAAACGTCGACTTGTGTTGGGGTTGAGGGAGGTTCTC
35 AAACACCTGAAGCTCAAAAACTGAAATGTGTCAATTATTTCTCCCAACTGTGAGAAGATACAGTCAAAAGTGGCTGGATGAC
ACTTTGCACACAATTTATGATTATGCTGTGAGCAGAACATTTCCCTTTGTGTTTGTCTCAACCGCAAAGCTCTGGGGCGCAGT
TTGAATAAGGCAGTTCTGTGATGTTGGTGGGATCTTCAGCTATGATGGGGCCAGGATCAGTTCCACAAGATGGTTGAGCTG
ACAGTGGCGGGCCGACAGGCGTACAAGACCATGTGAGGAATGTGCAGCAGGAGCTGGTGGGAGAGCCAGGCTCAGGCACCT
CCCAGCCTACTCACACAGGGCCCCAGCTGCCCTGCAGAAGATGGCCCCCAGCCCTGAAAGAAAAAGAGGCCACACTACATT
40 GAAATCTGGAATAACATCTGGAAGCATACAGTGGATGTACCTTGGAGCTAGAAGAATCCTTGGAGGCTTCAACCTCTCAAATG
ATGAATTTGAATTTATGAGAGTTCTTGCCCTGTGTCTGTATTTTGGGTAAAGAGGGGAGGCTGAAAAAGACTTTGGGGCTTT
TTCTTCTGTTTTTCATGACAAATGTAATTTGTGTAAGTGTGAAATCTGGAATTTGATCAGCATTAAGGGCAGATGAAGCAGTGT
CTGCAGGCGTTTCAGTGTGCGGAGCCTGTTAAAGGTCACTCAGATGTGAGGTGTTAATCTTCTTAAAGCCCTGGTGATACAG
CTCTGGCTTTCTGAGCACACTACGGATCTGGAATACTGGAATACTGATACTTAGAATACTTTGGCTGCTAAGGAACTTCC
45 TCTCCATTGCAGAATAGCTGAGCCAAGTGAGTGAGTTGAGCAAGCAGGTGGTGAGTCTCTGCTGTGGAGGTGGCATGGA
GGGCCATTCTGCCCGGCAACAGCACCGTCTGCAGGGAGCCACTTGGCAGAAGGGTGAGGGCTGCTGGTGTGAGAGCAAGAG
GGCTACAGGGAAGGGCCCTTTCTCAGGGGATGTAGCTTTTTTAAAGATTGGGAACACTTGGAGGATTTGCTAAATGAGCC
TCAGAAGGAAATTTGGTTTTCTAACCTGTGACTTTTTGAAATGAATTTATCTTTCAGTCTTTATTTTTCAAAGAAACAATGTG
TATTGAAGTACCTAGATTGTGTTGATAATCAACAACTTTCTCTTTTCAATGAACATATCTGAATGTGGTTTTCTGTCTTAGA
50 CCAGGAGGACAGAGTTTGTCTTTCATATTTTCCCTGTAGTAAGAGGGCTTATTTATTTTAAATAAGAGTAATTTATTA
AAAAA

63

64

CGCAGGAGGAAACGGTGCCGGAGCGCGCAGGGCTTGCTGCCGCCACCGCCGTGCACAGGCTGCCGGAGCGAGCCTGCCGCGCG
CCGCCCTCCCGCTCTCCTTCTCGGCGAGCTGCGGGGATGGGGCGGCCGGGAGCCCGAGCGCGCGCAGGAACCGCCGCCGC
CGCCGCCCGCTCTCCGTTGCGCGCGCGCTGAGCCGCGTCCGCCCGCGCGCGCTTCCCGGGGGCGGCCCGCCGAGCCCATG

GGCCGGAACAACCGCTGGTACCTGGCGGTGGCCGCCACCTACGTGCTGCCTGAGCCGGAGACGGCGAGCCGCTGCAACCCCGCG
GCATCCGACCACGACACGGCCATCGCGCTCAAGGACACGGAGGGGCGCAGCCTGGCCACGCAGGAGCTGGGGCGCCTCAAGCTG
TGCAGAGGGCGGGGCGAGCCTGCACTTCGTGGACGCTTTCTCTGGAACGGCAGCATCTACTTCCCTACTACCCCTACAACATAT
ACGAGCGGCGCTGCCACCGGCTGGCCAGCATGGCGGCATCGCGCAGAGCACCAGGTGCTGTTCAGGGCCAGGCATCCCTC
5 GACTGCGGCCACGGCCACCCGACGGCCGCCCTGCTCCTCTCCTCCAGCCTAGTGGAGGCCCTGGACGCTTGGGCGGGAGTG
TTCAGCGCGGCCGCTGGAGAGGGCCAGGAGCGGCGCTCCCCACCACCGCGCTCTGCCTCTTCAGAATGAGTGAGATCCAG
GCGCGGCCAAGAGGGTCAGCTGGGACTTCAAGACGGCCGAGAGCCACTGCAAGAAGGGGATCAACCTGAAAGAGTCAACCA
ATCGCATCATCTACCTTGATCCAATCCGACCTGACATCGTTTATGGCACCGTGTAATGAACAGGACTGTTTTATTCTTGGG
ACTGGAGATGGCCAGTTACTTAAGGTATTTCTTGGTGAGAAATTTGACTTCAAAATGTCAGAGGTATCTATGAAATTAAGAA
10 GAGACACCTGTTTTCTACAACTCGTTCCGTGATCCTGTGAAGAAATATCTACATTTATCTAACAGCTGGGAAAGAGGTGAGGAGA
ATTCGTGTTGCAAACTGCAATAAACATAAATCCTGTTCCGAGTGTTTAACAGCCACAGACCCCTACTGCGGTTGGTGCCATTG
CTACAAAGGTGCACTTTTCAAGGAGATTGTGTACATTACAGAACTTAGAAAACGGCTGGATATTTCTGCTGGAGCAAAAAAG
TGCCCTAAAATTCAGATAATTCGAAGCAGTAAAGAAAAGACTACAGTGACTATGGTGGGAAGCTTCTCTCCAAGACACTCAAAG
TGATGTTGAAGATGTGGACTCTAGCAGGGAGCTTGCAGAAATAAAGTCAAGCCAAACCGGACCTGCACCTGTAGCATCCCA
15 ACCAGAGCAACCTACAAGATGTTTCAGTTGTCAACGTGATGTTCTCTCTCGGTTCTTGAATTTATCAGACAGATTCAACTTT
ACCAACTGCTCATCATTAAAGAAATGCCAGCATGCGTAGAAACTGGCTGCGCGTGGTGTAAAGTGAAGAAGGTGTATCCAC
CCCTTCACAGCTTTCGACCTTCTGATTATGAGAGAAACAGGAACAGTGTCCAGTGGCTGTCGAGAAGACATCAGGAGGAGGA
AGACCAAGGAGAAACAGGGGAACAGAACCAACAGGCTTTACAGGCTTCTACATTAAGTCCATTGAGCCACAGAAAGTATCG
ACATTAGGGAAGCAACGTGATAGTAACGGGAGCAAACTTTACCCGGGCATCGAACATCAATGATCCTGAAAGGAACAGT
20 ACCTGTGATAAGGATGTGTACAGGTTAGCCATGTGCTAAATGACCCACATGAAATTTCTCTTCCATCAAGCCGGAAGAA
ATGAAGGATGTGTATCCAGTTTGATGGTGGGAAGTGTCTTCTGTGGGATCCTTATCCTACATTGCTCTGCCACATTGTTCC
CTTATATTTCTGTCTACCCTGGATCAGTGGTGGTCAAAATATAACCATGATGGGCAGAAATTTTGATGTAATTGACAACCTTA
ATCATTTACATGAATTAAGGAACATAAATGTCTCTGAATATTTGTGTGGCGACTTACTGCGGGTTTTAGCCCCCAGTTTA
AAGAGTTCAAAAGTGGCGACGAATGTCACTGTGAAGCTGAGAGTACAAGACACCTACTTGGATTGTGGAACCTGCAGTATCGG
25 GAGGACCCAGATTACGGGGTATCGGGTGAATCCGAGGTGGACACAGAACTGGAAGTGAATAATCAAAAAGAAATGACAAC
TTCAATATTTCAAAAAGACATTGAAATTTACTCTCTCCATGGGGAAATGGGCAATTAATTTGCAGTTTTGAAAATATTACT
AGAAATCAAGATCTTACCACCATCTTTGCAAAATTAAGGCAATCAAGACTGCAAGCACCATTGCCAATCTTTAAGAAAGTT
CGGGTCAAGCTGGGAAACCTGGAGCTCTACGTGAGCAGGAGTCAAGTTCCTTCCACATGGTATTTCTGATTGTGCTCCCTGTC
TTGCTAGTGATTGTCAATTTTTCGCGCCGTGGGGTGACAGGCACAAATCGAAGGAGCTGAGTCGCAACAGAGTCAACAACTA
30 GAATGTCTGGAAGCGAGCTCCGGAAGAGATACGTGACGGCTTTGCTGAGCTGCAGATGGATAAATGGATGTGGTTGATAGT
TTTGGAACTGTTCCCTTCTTGTACTACAAACATTTTGTCTGTGAGAACTTCTTCCCTGAGTCAGGTGGCTTACCCACATCTTC
ACTGAAGATATGCATAACAGAGACGCCAACGACAAGAATGAAAGTCTCACAGCTTTGGATGCCCTAATCTGTAATAAAGCTTT
CTTGTACTGTCTACCCACACCTTGAAGAGCAGAAGAACTTTCTGTGAAGGACAGGTGTCTGTTGCTCCTTCTTAACCATTT
GCAGTCAAAACCAAGCTGGTCTACCTGACCAGCATCTTAGAGGTGCTGACCAGGACTTGATGGAACAGTGTAGTAACATGCAG
35 CCGAAACTCATGCTGAGACGCACGGAGTCCGTGCTGCAAAAACCTCTCACAACGGATGTCGCTGCTGCTTTCTGGATTCTC
CGGGAGACTGTCCGAGAGCCCTTCTATTTGCTGGTGACGACTCTGAACAGAAAATTAACAAGGTCCTCGTGGATGTAATCACT
TGCAAAAGCCCTGTACACACTTAATGAAGACTGGCTGTTGTGGCAGGTTCGGAATTCAGTACTGTGGCATTAACCGTCTGCTTT
GAAAAAATCCCGGAAAACGAGAGTGCAGATGCTGTGCGGAATATTTCACTCAATGTTCTCGACTGTGACACCATTTGGCCAAAGCC
AAAGAAAAGATTTTCAAGCATTTTAAGCAAAAATGGCTCTCCTTATGGACTTCAGCTTAATGAAATTTGGTCTTGAGCTTCAA
40 ATGGGCACACGACAGAAAGAACTTCTGGACATCGACAGTTCCTCCGTGATTTCTGAAGATGGAATCAACCAAGCTAAACACCATTT
GGCCACTATGAGATATCAATGGATCCACTATAAAGTCTTTAAGAAGATAGCAAAATTTTACTTCAGATGTGGAGTACTCGGAT
GACCACTGCCATTTGATTTTACCAGATTCCGAAGCATTCAGAGATGTGCAAGGAAGAGACATCGAGGGAAGCACAAGTTCAAA
GTAAAAGAAATGTATCTGACAAAGCTGCTGTGACCAAGGTGGCAATTCATTCTGTGCTTGAAAACTTTTGAAGCATTTGG
AGTTTACCCAACAGCAGAGCTCCATTTGCTATAAAATACATTTTGTGACTTTTGGACGCCCAGGCTGAAAAAATAAATCACA
45 GATCCTGACGTCGTACATATTTGGAAGCAACAGCCTTCTCTTCTGCTGCTGGGTAAACATCCTGAAGAACCTCAGTTGTGTC
TTTGACATTAAGAAGACACCATATAGACGGCTGTTGTGTCAGTGATTGCCAGGCATTTCATGGATGCATTTTCTCTCACAGAG
CAGCAACTAGGGAAGGAAGCACAATAAAGCTTCTCTATGCCAAGGATATCCCAACCTACAAAGAAGAAAGTAAATCTTAT
TACAAAGCAATCAGGGATTGCTCTCATCTCAGAAATGGAAGAAATTTTAACTCAGGAATCTAAGAAACATGAAAT
GAATTTAATGAAGAAGTGGCTTGACAGAAATTTACAAATACATCGTAAAAATTTTGTATGAGATTCTAAATAAAGTAAAGAA
50 GAACGAGGGCTGGAAGAAGCTCAGAAACAACTCTTGCATGTAAAGTCTTATTTGATGAAAGAAAGAAATGCAAGTGGATGTAA
GCATCTGGGGCTGGCTTAATCTGGCAAGTCTTTCAGACGACTTGGGAGCAAAATGGCTGCTTGAGCTACTCTGTGCTGTTA
ATTTGTTGTTGACATAGGTTCCACTTTGGGCACTGTCTTTTTAAGAGACCAAGGCACATGCACAGCTTTTAGAAAGCAA

66

ACGCGTCCGAAGACATTAAGTAAAAAATGGAACATATGATTTTCTTTGTCATTTTAAAAAAGAATTATTTTATTAACCTGC
TGGCATATAATCTGGAGTCTTTTCACAACCTTACTTTTCTGATTGCTTTATGAATGATTGAATACATCTTCTTTCTAAA
AATATGTTGTAAATCTCCCTTGGCAAGATTTCTCCCTATGAGGGTAGTTATTTATTTGAGTCTGCCAAGTGGTTACCATGGGGC
5 AAGGTGCCATGATGTATTTCTTGGGTGCATTTGGTTTTCGCGCATGTAAATTTAAGACACTTATAGTAAGTGGACTCATTCTATA
GATGAGTTTCAGAACCTTTTACGTTCTCGGTAGAGGCTTCTGTCGGACAGGCAGAGAGTGTATTCCTCACTTTTCTTTTGTCT
TTCAAATTCAGTAAGGCATAGCACTTTTAAGAAATAGAATTTTCTATCATCTATGCAAATGATATTTATGTTAATATTA
TATCTTATGTTACACTGGGAGTAATTTGAGGTGCAATTTTATTACTACTTTGAATAGAGGACCATTATCTTCTTCTTCTTC
AGAAACTAAGAAGTAAGTGTAACTTTTAAAGTAAGTATATATCAGTGAGAGTAGGCTTGTTTTACAACATATTTCTAGCCAGTG
10 AGTTGTGTTTTCATGTCTCATCAAAAGACAATACCACATTGCATCATTTTACAAAATATGTTTGTCTATTTTCATTTCAGTTGTAA
CATAGGAAAATAGATATTTCTTAGATGATTTCTGAGTTTCTTACTGCAAAGAACAGTTATAAATGGTATACATGTGTCTCTGT
AATAGGGATAATATTGATATATCTGTGTCTACATATTTAAGAATCATCTATCTTATGTTGTCTTGAGGCCAAGATTTACCACG
TTTGGCCAGTGATTTGAATTTGGTGGTAGAAGGTAGTTCCATGTTCCATTTGTAGATCTTTAAGATTTTATCTTTTGATAACTTTA
ATAGAATGTGGCTCAGTTCTGGTCTTCAAGCCTGTATGTTTGGATTTTTCAGTAGGGACAGTTGATGTGGAGTCAATCTCTT
15 TGGTACACAGGAAGCTTTATAAAATTTCAATTCACGAATCTCTTATTTTGGGAAGCTGTTTGCATATGAGAAGAACTGTTGA
AATAAGGAATAAGCTTTATATATTGATCAAGGTGATCTGAAAGTTTAAATTTTAAATGTTGTAATGTTATGTTATGTTAA
TTGTACTTTATTTATGTATTCATAGAAAATCATGATTTATTAATAAAAGCTTAAATTCATCTAAAAAATAAAAAA

67

MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAQRFCQVSGYLDCTCDVETIDRFNNYRLFPRQLQKLLSDYFRYYKVNLRKRP
20 CFWNDISQCGRRDCAVKPCQSDVDPDGIKSASYKYSEBANNLIEECEQAERLGAVDESLSEETQKAVLQWKHDDSSDNFCEA
DDIQSPEAEYVDLLNPERYTGKGPDAWKIWNVIYEENCFKPTIKRPLNPLASGQGTSEENTFYSWLEGLCEVKRAFYRLIS
GLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLIELRALSQVLPFFERPDFQLFTGNKIQ
DEENKMLLEILHEIKSFPLHFDENSFFAGDKKEAHKLKEDFRLHFRNISRIMDCVGCFCRLWGKQLQTOGLTALKILFSEKL
IANMPESGPSYEFHLTRQEIIVSLFNAFGRISTSVKELENFRNLLQNIH

25 68

GCACGAGCCCCGGGCTGCCGGCGCGGGCGCCGCGCACGCTCCACAGGCTGGGTCCGCGAGGTGGCGATCGCTGAGAGGCAGGAGG
GCCGAGGCGGGCTGGGAGGCGGCGCGGAGGTGGGGCGCGCTGGGGCGCGCCCGCACGGGCTTCATCTGAGGGCGCACGGCCC
GCGACCGAGCGTGGGACTGGCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCAATGGGCGCGGCTGGGGATTCTTGT
TGGCTCTCTGGGCGCGGTGTGGCTGCTCAGCTCGGGCCAGGAGGAGCAGCCCCGAGACAGCGGCACAGAGGTGCTTCTG
30 CCAGGTTAGTGGTTACTTGGATGATTGTACCTGTGATGTTGAAACCATGATAGATTTAATACTACAGGCTTTTCCCAAGACT
ACAAAACTTCTTGAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAGCCGCTGCTCTTCTGGAATGACATCAGCCA
GTGTGGAAGAAGGACTGTGCTGTCAAAACCATGTCAATCTGATGAAGTTCTGATGGAATTAATCTGCGAGCTACAAGTATTC
TGAAGAAGCCAATAATCTCATTGAAGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTCTGAGTGAGGAACACA
GAAGGCTGTTCTTCAGTGGACCAAGCATGATGATTTCTTCAGATAACTTCTGTGAAGCTGATGACATTACGTCCCTGAAGCTGA
35 ATATGTAGATTTGCTTCTTAATCTGAGCGCTACACTGGTTACAAGGGACCAGATGCTTGGAAAATATGGAATGTCATCTACGA
AGAAAATGTTTAAAGCCACAGACAATTAAGACCTTTAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGAACTTT
TTACAGTTGGCTAGAAGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAATGTGCA
TTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTGA
TGGAAATTTGACTGAAGGAGAAGGTCCAAGAAGGCTTAAGAATCTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAA
40 AGTGTTACCATTCTTCGAGCGCCAGATTTTCAACTCTTTACTGGAATAAAATTCAGGATGAGGAAACAAATGTTACTTCT
GGAAATACTTCATGAATCAAGTCATTTCTTTGCAATTTGATGAGAATTCATTTTCTGCTGGGATAAAAAGAGCACACAA
ACTAAAGGAGGACTTTGACTGCAATTTAGAAAATTTTCAAGAATTATGGATTGTGTTGGTTGTTTAAATGTCGTCTGTGGGG
AAAGCTTCAGACTCAGGGTTTGGGCATGCTCTGAAGATCTTATTTTCTGAGAAATGATAGCAAAATATGCCAGAAAGTGGACC
TAGTTATGAATTCATCTAACCAGACAAGAAATAGTATCATTTTCAACGCATTTGGAAGAATTTCTACAAGTGGAAGAATTT
45 AGAAAATTCAGGAATTTGTTACAGAATATTCATTAAGAAAACAAGCTGATATGTGCCTGTTTCTGGACAATGGAGGCGAAAG
AGTGGAAATTTCAATCAAGGCATAATAGCAATGACAGCTTAAAGCCAAACATTTTATATAAAGTTGCTTTTGTAAAGGAGAAT
ATATTGTTTAAAGTAAACATTTTAAAAATCTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAGTAATACTTTAATAA
TGTGCTACAAATTTAAAGTTAATATGAATAAAAGGAGGATTATCAAAATTCATATGATAAAAGTGAATGTTCTAAGTCTC
TCAAACTAGCGTTTATGTAATAATATGAATAAAATAAACTATGGTAAATGTGACAAGCATTTAATAGGAAATGCTAAGG
50 AGGCCTCATAAATGACCATAATTACCAACGTAGAATTTTTCAGTACATTTAGGCTTGTGGAATTTAGCAAAATAAAAATAAGA

TTGCCAGTTAGATTTGAATTCAGATAAAACAATTAGTTTTTTAATATTTTACATGGAATATTTGAAAAACTTATACTAAAA
AATTATTTGTTTGAATTCACATTTAACTGGGAGTCTTGTATTTTATCTGGCAATCCTAAAAATACATTTGGTATGAAACAAATCA
CTTTTAGAAGTATATTGCTATTTTGATTGGGTGTTTTTGTGTGTAGAAACGTACAATAACAACCTCAAAGGCACAGGAGATTTC
TAAACATTTGTGAAAAGTTGAATAGATTATATATTTATTTCTCATAATACTTTCTACTAATACTAAATAAAATTTGGGGAACACTTT
5 TTATTTTATATAATTTCCAATTTACAGAAAAGTTTCAAAAATAGTACAAAGAGCTCTCTTACCAGATTCTACTAATTTGTTTCAT
ACGTGCTTTATCTTTTCATGCTTTCTCTGTACACACACACACACACAAATTTTCCCTCAATCATTTGAAAGTCAGTTATAGGC
ATCATGCCCCCTAAACCCCTAAATACTTCAGTGTGTAATACTGAATAATTACTAAAAATGATTTTCTCAGAAAAAAAACCTCCCA
CAATTCGGAACATAATACTGTAAAGCCTTAGAATAAATAACTTTTCAAGTTCCAATCTAAAGTTCTTTTGTAGTTTGTGTGC
CCGTTTTATGCTTGATGTGTATAGTAATAGGGTAGGCTATTTATTTTATTTAAATTTTTTTAGAGACAAGGTTTGTCTGTGTT
10 GCGCAAGCTGGAACCTGAACGACTGGGCTGAAGTGACTTCCACCTCAGCCTCCCAAGTAGCTGGGAATACAGGTGTCTGCCA
CCATACCCAGTTTCAATTTTGTTTTATACCCGAAGTTTCATTTCCCTTGTCTCCCTAAAACTGAAGCTTAATTTTGGGAGGTT
TTCATTAGTGGAAGCTCTTCATTTATAAAGCTATTTGAAGGGGTTTAGGAATTTATATCACATGGTAATTTGTAGAGAAAAAGAA
GCTATATACCTCAAAATCGTGCCCTCTTTACATATGTCTTATCAGGTATAACATGTTGAAATGTACATTAGTAGTAAAGTGGG
GTTTATTTATATAGTGGTTAAGAAATGTAGTTTACACTGCTGTATACCTTCTTCTGTGTCCCTAAGGCCTGGTACAGTGCC
15 AAGCACATACTTGGTATCCAATAAATATTTGTTGGATGAAAAAAAAAAAAAAAAAAAAA

70

GGCAGGAGGATGGCGGCAGCGCGCTGGGAGGGCGAGGCGGAGGCGGCAAAACGGGCGGTGAGCAGAACGTGTAGCCGCG
TCCCTTCCAGTCCGCTCCGGGCAGCTGCTGATGCAAGGAATCCCTTGGGCTCCCGTCCACTCCACTGCTGACCAGCCCATTCGC
CTGTGCTGAGTCTTCCCTGCAGGCCCTTCCCTTGCCCTCTGTGGGACCCCTGTGGGGTCCATCCGGCTGGAGAAGAAAGCCTCTCA
20 TGCTAACGTTGCAGACCCAGAGGGTCCCTGTGTGGGTGTGGAGATGGCCAATGAGAATCACGGCAGCCCCGGGAGGAAGCGTC
CCTGCTGAGTCACTCCCCAGGTACCTCCAATCAGAGCCAGCCCTGTTCTCCAAAGCCAATCCGCTGGTTCAGGACCTCCCAGA
GGAGCTGGTGCATGCAGGCTGGGAGAAGTGCTGGAGCCGGAGGAGAATCGTCCCTACTACTTCAACCGATTACCAACCAAGTC
CCTGTGGGAGATGCCCCGTGCTGGGGCAGCAGATGTGATTTTCGGACCCCTTGGGGTGAATGCGACCCCACTGCCCCAAGACTC
AAGCTTGGTGGAACTCCCCGGCTGAGAACAAGCCAGAAAGCGGAGCTCTCGGAAGAGCAGCAAGCGGCAATGGTGTGAA
25 GAAGCCCAAGATTGAAATCCAGTGACACCCACAGGCCAGTCCGTGCCAGCTCCCCAGTATCCAGGAACCCCAACGCTGAA
GATGTGGGTACGTCCCCGTAAGATAAACACAGGCAGCTCTCTACGACCCACTGAGGTCTACTGGGACCTGGACATCCAGAC
CAATGCTGTCAATCAAGCACCGGGGGCCTTCAGAGGTGCTGCCCGCATCCCGAAGTGGAATGCTCCGCTCTCAGCTCATCCCT
GAAGCTTCGGCAGCACTATCGGGAGCTGTGCCAGCAGCGAGAGGCAATGAGCCTCCACGGGAGTCTTTCAACCGCTGGATGCT
GGAGCGCAAGGTGGTAGACAAAGGATCTGACCCCTGTTGCCAGCAACTGTGAACCAAGTCGTGTCACCTTCCATGTTTCGTGA
30 AATCATGAACGACATTCCTATCAGGTTATCCCGAATCAAGTTCCGGGAGGAAGCCAAGCGCCTGCTCTTTAAATATGCGGAGGC
CGCCAGGCGGCTCATCGAGTCCAGGAGTGATCCCTGACAGTAGGAAGGTGGTCAATGGAATGTGGAAGACACCTTTAGCTG
GCTTCGGAAGGACCACTCAGCCTCCAAGGAGGACTACATGGATCGCCTGGAGCATCTGCGGAGGAGTGTGGCCCCACGCTCTC
GGCCGAGCCAAGGACTCCGTGGAAGGCATCTGCAGTAAGATCTACCACATCTCCCTGGAGTACGTCAAACGGATCCGAGAGAA
GCACCTTGCCATCCTCAAGGAAAACAACATCTCAGAGGAGGTGGAGGCCCTGAGGTGGAGCCCCGCTAGTGTACTGTACCC
35 AGTCCGGCTGGCTGTGTCTGCACCGCCATGCCAGCGTGGAGATGCACATGGAGAACAACGTGGTCTGCATCCGGTATAAGGG
AGAGATGGTCAAGGTCAAGCCGCAACTACTTCAGCAAGCTGTGGCTCCTTTACCGCTACAGCTGCATTGATGACTCTGCCTTTGA
GAGGTCTCTGCCCCGGGTCTGGTGTCTTCTCCGACGGTACCAGATGATGTTCCGGCTGGGCTCTACGAGGGGACTGGCCTGCA
GGGATCGCTGCCTGTGCATGTCTTTGAGGCCCCCACCAGCTCTTTGGCGTCAGCTTCGAGTGTCTCGCCTCACCCCTCAACTG
CTACTTCCGCCAGTACTGTCTGCTTCCCGACACAGACGGCTACTTTGGCTCCCGGGGCCCTGCCTAGACTTTGTCTCACT
40 GAGTGGTTTCAATTTGAGGCCAACCTCCCTTCTGCGAGGAGCTCATGGATGCCATGGTCTCTCACTTTGAGAGACTGCTTGAGAG
CTCACCGGAGCCCCGTCTTCATCGTGTTCATCCCTGAGTGGCGGGAACCCCAACACCAGCGCTCACCCGCATGGAGCAGAG
CCGCTTCAAACGCCACCAAGTTGATCCTGCCCTGCCCTTTGAGCATGAGTACCGCAGTGGCTCCAGCACATCTGCAAGAAGGAGGA
AATGCACATACAAGGCGCTCCACAACACGGCTGTGCTCTTCCCTACAGAACGACCTGGCTTTGCCAAGTGGCGCCGACGCTGA
ACGGCTGCAGGAGCTGAGTGTCTGCCATACCGGCAGTGCAGGCCAGCCACAGCTCTGGTTCTTCTCTCATCTCTCTCGGAGGC
45 CAAGGACCGGGAGCTCGGGCGTGAGCAGGGTCTTAGCCGCGAGCCTCACCCCACTTAACATATCTCGGGGAGGAGGAGCCCC
AGGGGTGCTAGTCTGGACTGCTGGGACTCGGGCCCCGCGGCTCAGAGGGACCCGGCTGCCACTGACATATGAAGATTATGG
TTCTGCCAGGGCTCCCCCTCCCTGCTGTCCCAAGTCTCACCTCAAACCTCCCTCAAGTCCCATGTATATAGGTCTGTATGCC
TTCCCAACCCCGCCCTCACCTGTGTCCACCTTGTTCATTTGTAAAAGGAAATACAGAAACCCCCCAAAAAAAAAAAAAA
AAAAAAAAAAAA

50 72

[illegible]

73

25 MREIVHIQAGQCGNQIGAKFWEVISDEHGIDPSGNYVGSDQLQLERISVYNEASSHKYVPRAILVDLEPGTMDSVRSAGFGL
FRPDNFI FGQSGAGNNWAKGHYTEGAELVDSVLDVVRKECENCDC LQGFQLTHSLGGGTGSGMGTLLISKVREEYPDRIMNTFS
VVPSPKVS DTVPEYPNATLSIHLVENTDBTYCIDNEALYDICTFTLKLATPTYGDNLHLVSATMSGVITSLRFPGLNADLRK
LAVNMVFPRLHLFVFKFAPLTRRGSQQYRALTVPELTQOMFDAKNMMAACDPRHGRYLTVATVFRGRMSMKVEDEQMLAIQSK
NSSYFVWEI PNNVWVAGDIPRGLKMSSTFIGNSTAIQELFKRISEQFTAMFRKRAFLHWYTGEGMDEMEFTEAESNMNDLVS
30 EYQYQYQDATAEEGEMYEDDEEESAQSPK

74

ATGCGGGAGATCGTGACATCCAGGCCGGCCAGTCGCGCAACAGATCGGGGCCAAGTTCCTGGGAAGTCATCAGTGATGAGCAT
GGCATCGAACCCAGCGGCAACTACGTGGGCGACTCGGACTTGCAGCTGGAGCGGATCAGCGTCTACTACAACAGAGCCTCTTCT
CACAAGTACGTGCCCTCGAGCCATTCTGGTGGACCTGGAACCCGGAACCATGGACAGTGTCCGCTCAGGGGCCCTTTGGACATCTC
TTCAGGCCCTGACAATTTTCATCTTTGGTCAGAGTGGGGCCGGCAACAACCTGGGCCAAGGGTCACTACACGGAGGGGGCGGAGCTG
GTGGATTCCGTCTCGGATGTGGTGCAGGAAGGAGTGTGAAACTGCGACTGCCTGCAGGGCTTCCAGCTGACCCACTCGCTGGGG
GGGGGACGGGCTCCGGCATGGGCACGTTGCTCATCAGCAAGGTGCGTGAGGAGTATCCCGACCGCATCATGAACACCTTCAGC
GTCGTGCCCTCACCAAGGTGTGAGACACGGTGGTGGAAACCTTACAACGCCACGCTGTCCATCCACCAGCTGGTGGAAAAACAG
GATGAAACCTACTGCATCGACAACGAGGCCTCTACGACATCTGTCTCCCGACCTTCAAGCTGGCCACGCCCACCTACGGGGAC
CTCAACCACCTGGTATCGGCCACCATGAGCGGAGTCAACCACCTCCTTGCGCTTCCCGGGCCAGCTCAACGCTGACCTGCGCAAG
CTGGCCGTCAACATGGTGCCTTCCCGCGCTGCACTTCTTCATGCCCGGCTTCCGCCCCCTCACCCAGCGGGGCGGCCAGCAGCAG
TACCGGGCCCTGACCGTGCCCGAGCTCACCCAGCAGATGTTTCATGCCAAAGAATATGATGGCCGCTTGCACCCGCGCCACGGC
CGCTACCTGACCGTGGCCACCGTGTTCCGGGGCCCGCATGTCCATGAAGGAGGTGGACGAGCAGATGCTGGCCATCCAGAGCAAG
AACAGCAGCTACTTCGTGGAGTGGATCCCCAACACGTGAAGGTGGCCGTGTGTGACATCCCGCCCCGCGGCCCTCAAGATGTCC
TCCACCTTCATCGGGAACAGCACGGCCATCCAGGAGCTGTTCAAGCGCATCTCCGAGCAGTTACGGCCATGTTCCGGCGCAAG
GCCTTCCTGCACATGGTACACCGGCGAGGCGATGGACGAGATGCGGACTTCCACCGAGCCGAGAGCAAGATGAGCAGCTGGTGTCC
GAGTACCAAGCAGTACCAGGACGCCACGGCCGAGGAAGAGGCGAGATGTACCAAGACGACGAGAGGAGTGCAGGCGCCAGGGC
CCCAAGTGAAACTGCTCGCAGCTGGAGTGAGAGGACGGTGGCGGGCCGGGCCGAAGCCAGCAGTGTCTAAACCCCGGAGCCAT
CTTGCTGCCGACACCCTGCTTCCCATCGCCCTAGGGCTCCCTTGCCCGCCCTCCTGCAGTATTTATGGCCTCGTCTCCCCA

CCTAGGCCACGTGTGAGCTGCTCCTGTCTGTCTTATTCAGCTCCAGGCCTGACGTTTTACGGTTTTGTTTTTACTGGTTT
GTGTTTTATATTTTCGGGGATACCTAATAAATCTATTGCTGTGACATACCCCT

75

5 MAEGETESPGPKKCPYISSVTSQSVNLMIRGVVLFFIGVFLALVLNLLQIQNVTLFPPDVIASIFSSAWVPPCCGTASAVI
GLLYPCIDRHLGEPHKFKREWSSVMRCVAVFVGINHASAKVDFDNNIQLSLTLAALSIGLWWTFRSRSRSGFGLGVGIAFLATVV
TQLLVYNGVYQYTSDFLYVRSWLPICFFAGGITMGNIGRQLAMYECKVIAENLIRNEEGKKYLLYRKAR

76

10 ACTCACTATAGGGCTCGAGCGGCCGCCGGGCAGGTGGGATTTTCGGGTAGGTCTACTTTAGGACAAGATGTGGTACCGTTGAA
GCGTCAGTCTTTGATTCACAGACAGTTGAGCTTTTCAGCTGGGAAGCCTTTCCATTTTTTTTTTTTAAACGGCTTTCTGAACCT
ATGAAACCATGGCAGAAGGAGAGACAGAGTCACCTGGGCCCAAAAAGTGTGGCCCATATATTTTCATCTGTCACTAGCCAGAGTG
TGAACCTTGATGATTCGAGGAGTAGTGCTATTTTTTATTGGAGTATTTCTTGCATTAGTGTTAAATTTACTTCAGATTCAGAGAA
ATGTGACGCTCTTTCCACCTGATGTGATTGCAAGCATCTTTCTTCTGCATGGTGGGTACCCCATGCTGTGGCAGCGCTTCAG
CTGTGATTGGGTATTATACCCCTGCATTGACAGACATCTAGGAGAACCACATAAATTTAAAAGAGAGTGGTCCAGTGAATGC
GGTGTGTAGCAGTCTTTGTTGGTATAAATCATGCCAGTGCTAAAGTGGATTTTCGATAACAACATACAGTTGTCTCTCAGACTGG
15 CTGCACTATCCATTGGACTGTGGTGGACTTTTGATAGATCTAGAAGTGGTTTTGGCCTTGGAGTAGGAATTGCCTCTTGGCAA
CTGTGGTCACTCAACTGCTAGTATATAATGGTGTTTACCAATATACATCTCCAGATTTCTCTATGTTCGTTCTTGGTTACCAT
GTATATTTTTTGTGCGAGGCATAACAATGGGAACATTGGTCGACAACCTGGCAATGTACGAATGTAAAGTTATCGCAGAAAATC
TCATCAGGAATGAAGAAGGCAAAAATATCTTTTGTACAGAAAAGCAAGATGAAAAGGATGTGAAATGGTAGATATACCAACAA
AACTTCAGACTGTAAAATTCAGGATGCAGTTTCCCTTGAATGGCGTGTGTGTATATATGGATAAATATATATATACACAC
20 ACACATATTAAGCAATCTGTGATTGCTTCATCTGTAATCAGTTGTAAACCTTTACATATTTGACTTAAATAACTGTAAGATA
TATATGTACTACATTAAGAGTGTGATTAATAGATGAAATTTTAAATTAATTTTTTAAACATGCCATACATTTGTATCACA
TGTTAATGTGCAAGATATTTCTCTCATGACAGATTAAGAATGCTTTGAACAATTTGTAGGACTTAGTGGAAATAAATA
AGAGGGAAGGCCAAAACAAACAAACAAAGGCATATGGGAGCTGGGTATTTCTCTTTAGGCTTACTGTTGTGGCTTTTTT
ATTTTTCTAATCCACGCCGGTATGGGGTTTGGGGGCCCAATTGTG

25 77

MASKLLRAVILGPPGSGKGTVCQRIQNFGLQHLSSGHFLRENKASTEVGEMAKQYIEKSLVDPHVITRLMSELENRRGQH
WLLDGFPRTLGQAEALDKICEVDLVISLNI PFETLKDRLSRRWIHPPSGRVYNLDFNPPHVHGIDVDTGEPLVQQEDDKPEAVA
ARLRQYKDVAKPVIELYKSRGVLHQFSGTETNKIWPYVYTLFSNKITPIQSKAY

78

30 CGGCGCTGGGCTGAGGGGAGGGTTGTCTTAAAGTCTCTCTTCCCTCTAGGGGCGGCCGCGAGTCCCAGTGAGAGCGGA
GGGTGCCAGAGGTAGGGGGCCGAGAAACAAAGTTCCCGGGCTTCTCCGGGGCCGCGTCCGGGCTGCGCGTTTGACCGCCCC
CCTCTCGCGAAGCAATGGCTTCCAACTCCTGCGCGCGGTCACTCTCGGGCCGCCCGCTCGGGCAAGGGCACCCTGTGCCAG
AGGATCGCCGAGAATTTGGTCTCCAGCATCTCTCCAGCGGCCACTTCTGCGGGAGAACATCAAGGCCAGCACCGAAGTTGGT
GAGATGGCAAAGCAGTATATAGAGAAAAGTCTTTTGGTTCCAGACCATGTGATCACACGCCCTAATGATGTCGAGTTGGAGAAC
35 AGGCGTGGACAGCACTGGCTCCTTGATGGTTTTCTTAGGACATTAGGACAAGCCGAAGCCCTGGACAAAATCTGTGAAGTGGAT
CTAGTGATCAGTTTGAATATTCATTTGAAACACTTAAAGATCGTCTCAGCCGCCGTTGGATTACCCCTCCTAGCGGAAGGGTA
TATAACCTGGACTTCAATCCACCTCATGTACATGGTATTGATGACGTCACTGGTGAACCGTTAGTCCAGCAGGAGGATGATAAA
CCCGAAGCAGTTGCTGCCAGGCTAAGACAGTACAAGACGTGGCAAAGCCAGTCATTGAATTATACAAGAGCCGAGGAGTGCTC
CACCATTTCGGAACGGAGACGAACAAATCTGGCCCTACGTTTACACACTTTTCTCAAACAAGATCACACCTATTTCAGTCC
40 AAAGAAGCATATTGACCCTGCCCAATGGAAGAACCAGGAAGATGTGGTCATTCAATAGTGTGTAGTATTGGTGCTGTG
TCCAAATTAAGAAGCTAGCTGAGGTAGCTTCAGCATCTTTCTAGTTGAAATGGTGAACATGATAGGAAAACAAATGAGTAGAAA
GAGTTTCATGAAGAGGCCCTCTCTGCCTTTCAAAGGCTGGTCACTACACATGTTAAGGTGTCTCTGCACATGTCTCAAGCC
CATCACAAGAAAGCAAGTACAGTGTGGATTTCAATGGTGTGTAACCTCAGCTCCAGCTGGTTTTTGACAGCTGTGTGTGGT
AATATTTTTGACATGTGATGGTGATAGTCTCTGGTTCTCCCATCCCCACAAAGGCTGTGAACCACAGCACCAGGAAGCCTGA
45 GAATGAATCCTGAGGGCTCTAGCCAGGCTTTGTCCAGGCTTTCTGGTGTGTGCCCTCCTGGTAACAGTGAATTTGAAGCTAC
TTACTCATAGTGGTTGTTTTCTCTGGTCTTGAGTGAAGTGTGTCCACAGTTTCAATTTTTTCCGGTAGGAATAACTCTTTTCTACA
TCCACGCTCCATAGAGTCTCTCTTTTCAGACATCTGGGATGAAAGAATTTGGCTTTTTTTTTTCTTTTTTTTGGACATCT
GTTTTCACTCTTAGGCTTTTAAACAATAGTTATTGCTTTTATCCCTCTCAGATTCTAATAACTGAGAGCGATGGGGCTATATTG

AATCTCTGTATGCACTGAGAACTGAGCTATGAAGAGAATCTTATTAAACTGCTGGTCTGACTTTATGGATTGACACTGTTCTCTT
TCTTTTATTGTGAAAAAAAAAAAAA

79

MGSVAAASAGVPRGCDILIVYSPDAEEWCQYLQTLFLSSRQVRSQKILTHRLGPEASFSAEDLSLFLSTRCVVLLSABLQVHFH
5 KPSLLPLLQRAFHPHRVVRLLCGVRDSEEFDFPDWAHWQELTCDDEPETYVAAVKKAISEDSCDSVTDTEPEDEKVVSY
KQQLPTVTSFPGNLMVVQPDRIIRCAETTVYVIVRCKLDDRVTGEAEFSPEDSPSVRMEAKVENEYTI SVKAPHLSSGNVSLKI
YSGDLVVCETVISYYTDMEEIGNLLSNAANPVEFMCQAFKIVPYNTETLDKLLTESLKNNIPASGLHLFGINQLEEDMMTNQR
DEELPTLLHFAAKYGLKNLTALLTCTPGALQAYSVANKHGHYPNTIAEKHGFRLRQFIDEYVETVDMKSHIKEELMHGEEAD
AVYESMAHLSTDLLMKCSLNPCCDEDLYESMAAFVPAATEELYVEMLOASTSNPIPGDGFSRATKDSMIRKFLGNSMGMTNLE
10 RDQCHLGQEEDEVYHTVDDDEAFSVDLASRPPVPVPRPETTAPGAHQLPDNEPYIFKVFAEKSQERPGNFVVSSESIRKGPVVRP
WRDRPQSSIIDPFAGMKT PGQRQLITLQEQVKLGIVNVDEAVLHFKEWQLNQKRSESFRRQENLKRRLDSITRRQREKQKSG
KQTDLEITVPIRHSQHLPAKVEFGVYESGPRKSVIPRTELRRGDWKT DSTSSTASSTSNRSSTRSLLSVSSGMEGDNEDNEVP
EVTRSRSPGPPQVDGTPPTMSLERPPRVPPRAASQRPPTRETFHPPPPVPPRGR

80

15 AGACCCCCAGGGTGCTCCGAGAGCTGCCTCACAGAGGACTCCGACCAGGGAGACCTTCCATCCTCCTCCACCTGTTCCACCC
AGAGGACGCTGATTCCACCTCCTAAAACCTGCCTACTTCAGGACTTTAAGACTCAGCTTTCAGCCTGTTAATGATGTCTTCA
TGTTGAGTTTATAGCATGACTGTTGACCTTAAGATCCATTCTCATTTGCTGATAATGCTGCAGCCCTGCTGGTTTGGGCTTGCT
CGAAGATTTTATTAAGGCACGAAGAAGTGAAAACTAAGGGCTTCATTACCATCACCAAGTATATCGAACCATATACTTGT
GCCAAAAGGGTGGAAGGACTTTAATTCGGAATACTTTACCTCTTAATTTGGCCATATTCAGGAGGCCATAAAGATTGATCCN
20 TAATGTTATTTACCGTGTTTATTGAATGG

82

TTTGTTGGTCTTTTTTTTTTTTCTCAGTGTTTCAATTCCTCCTTGGTTGAATCCATGGATGCAAAACCCACAGATATGAAG
GGCTGGCTATATATGCATTGATGATTGTCTATTATATTAGTTATAAAGTGTCATTTAATATGTAGTGAAAGTTATGGTACAGT
GGAAAGAGTAGTTGAAAACATAAACATTTGGGCCTTTCAAGAAAGGTAGCTTGGTGAAGTTTTCACCTTCAAACATATGTC
25 GTCAGGG

84

GCTTGAGATCTCCAGCCTTACCGCGGCTCGAAATGGCCCCAACTGCTCCTGCACCACTGGTGTCTCCTGCANCTGCACCGGCTC
CTGCAAGTGCAAGAGTGCAAAATGACCTCCTGCAAGAAGAGCTGCTGNNCTGCTGCCCCGTGGTGTGCAAGTGTGCCACGGTGT
GTCTGCAAGGGGACGTTGGAGAACTGCAGCTGCTGTNCTGATGTGGGAACAAGCTCTTCTCCAGATGTTAATAGAACAAAGCTG
30 CACAACCTGGGATTTTTTCAATACGATACTTGAGCCATTTCTGTCATTTCTTTTATTGTTAAATATGTGAGTGACAATAAA
ACAATTTGACTTTGAATCTTAAAAA

85

MPLGHIMRLDEKIALEYIVPCLHEVGFCYLDNFLGEVVGDCVLERVKQLHCTGALRDGQLAGPRAGVSKRHLRGDQITWIGGN
EEGCEAISFLLSLIDRLVLYCGSRLGKYVYKERSKAMVACYPGNGTGYVRHVDNPNGDGRCTICIYYLNKNWDALHGGILRIP
35 PEGKSEFIADVEPIFDRLFFWSDRRNPHEVQPSYATRYAMTVWYFDAEERAEAKKKFRNLTRKTESALTED

86

GAGTCTGGCCGAGTCGCGGCAGTGGTGGCTTCCCATCCCCAAAAGGCGCCCTCCGACTCCTTGCGCCGCACTGCTCGCCGGG
CAGTCCGGAACGGGTCGTGGAGCTCCGCACCACTCCCGCTGGTTCCCGAAGGCAGATCCCTTCTCCCGAGAGTTGCGAGAAAC
TTTCCCTTGTCCCCGACGCTGCAGCGGCTCGGGTACCGTGGCAGCCGAGGTTTCTGAACCCCGGGCCACGCTCCCGCGCCTC
40 GGCTTCGCGCTCGTGTAGATCGTTCCCTCTCTGGTTGCACGCTGGGGATCCCGACCTCGATTCTGCGGGCAGATGCCCTTG
GACACATCATGAGGCTGGACCTGGAGAAAATGCGCTGGAGTACATCGTGCCTGTCTGACAGAGGTGGGCTTCTGCTACCTGG
ACAACCTTCTGGGCGAGGTGGTGGGCGACTGCGTCTGGAGCGCGTCAAGCAGCTGCATGCACCGGGGCCCTGCGGGACGGCC
AGCTGGCGGGCGCGCGCGCGCTCTCAGCGACACCTGCGGGCGACCAAGATCAGTGGATCGGGGCAACGAGGAGGGCT
GCGAGGCCATCAGCTTCTCTCTGCTCATCGACAGGCTGGTCTCTACTGCGGGAGCCGGCTGGGCAAACTACGTCGAAGG
45 AGAGGTCTAAGGCAATGGTGGCTTGCTATCCGGGAAATGGAACAGGTTATGTTCCGCCACGTGGACAACCCCAACGGTGATGGT
GCTGCATCACCTGCATCTACTATCTGAACAAGAATTGGGATGCCAAGCTACATGGTGGGATCTGCGGATATTTCCAGAGGGGA

[illegible]

- GCCTCCCCAGGCCCGCGCCCTGAGCCAGGTTGCCATTGCCCGCACAGGCCCTATTCTCTCAGCCCTCGGCGGCGATGAGGC
GCTGAGGCGGCTGCCCGCGCTGCCCGGAGCTTAGGACTCGGAAGCGGCGGGCCGAGGCGTGCGGCGCTCCCTGAGG
CGAGGGTAGCGGTGCATGGCGCAGTAACGGCCCTATCTCTCTCCCGCTCCCCAGCCTCGGGCGAGGCCGTCCGGCCGCTAC
CCCTCTGCTCGGCCCGCGAGTCGCCGTGCCCGCCCGCCCGCCCGCATGGCCAATGACAGCGGCGGGCCCGCGGCCGAGC
5 CCGAGCGAGCGAGACCGGCAGTACTGCGAGCTGTGCGGGAAGATGGAGAACCTGCTGCGCTGCAGCCGCTGCCCGAGCTCCTTC
TACTGCTGCAAGGAGCACCAGCGTCAGGACTGGAAGAAGCACAAAGCTCGTGTGCCAGGCGAGCGAGGCGCCCTCGGCCACGGA
GTGGGCCACACACAGCATTCCGGCCCCCGCGCCCGCGCTGCAGTGCCCGCCCGCAGGCGGGGCGCGGAGCCAGGAAGGCA
GCGGCGCGCCGGGACACCGCTCCGGGGACGCGGCCAAGGGAAAAGTAAAGGCCAAGCCCCCGCGCCGACCCAGCGCGCGCCGCG
TCGCGTGTGCTGCGGCCCGCGCGCCAGGGCTCGGCGGTGGCTGCCGAAGCCGAGCCCGCAAGGAGGAGCCGCGCGCCGCG
10 TCATCGCTGTTCCAGGAGAAGGCGAACCTGTACCCCCAAGCAACACGCCCGGGATGCGCTGAGCCCCGCGCGCGCTGCGG
CCCAACGGGCGAGACGAAGCCCTGCCCGCGCTGAAGCTGGCGCTCGAGTACATCGTGCCTGCATGAACAAGCACGGCATCTGT
GTGTGGACGACTTCTTCGGCAAGGAGACCGGACAGCAGATCGGCGACGAGGTGCGCGCCCTGCACGACACCGGGAAGTTCACG
GACGGGCGAGCTGGTCAGCCAGAAGAGTGACTCGTCCAAGGACATCCGAGGCGATAAGATCACTGGATCGAGGCAAGGAGCCC
GGCTGCGAAACCATTTGGCTGCTCATGAGCAGCATGGACGACCTGATACGCCACTGTAACGGGAAGCTGGGCGACTACAAAATC
15 AATGGCCGAGCAAAAGCCATGGTGTGTTATCCGGGCAATGGAACGGGTATGTACGTCATGTTGATAATCCAAATGGAGAT
GGAAGATGTGTGACATGTATATATATCTTAATAAAGACTGGGATGCCAAGGTAAGTGGAGGTATACTTCGAATTTTCCAGAA
GGCAAGCCAGTTTGGCTGACATTGAACCCAAATTTGATAGACTGCTGTTTTCTGCTGACCGTCGCAACCTCATGAAGTA
CAACCAGCATATGTACAAGGTACGCAATAACTGTTTGGTATTTTGTATGATGAGAGACGAGCTAAAGTAAATATCTTA
ACAGGTGAAAAAGGTGTGAGGTTGAACCAATAAACCTTCAGATTGCGTGGTAAAGACGCTCTCTAGAGCCTTTGATCCAGC
20 AATACCCCACTTCACTACAATAATGTTAACTATTTGTTAACTGTGAAATACGAATAAATGGGATAAAGAAAAATAGACAACCA
GTTTCGATTTTAATAAGGAAACAGAAACAACCTTTTGTGTGTCATCAAAACAGAAAGATTTTGTACTGCTGTGACTTTGTACTGCAT
GATCAACTTCAAACTCTGTGATTGCTTACAGGAGGAAGATAAGCTACTAATTGAAAAATGGTTTTTACATCTGGATATGAAATAAG
TGCCCTGTGTAGAAATTTTTTTCATCTTATATTTTGCCAGATCTGTTATCTAGCTGAGTTTCACTCTCTCCCTTTTTTATA
TCAAGTTTGAATTTGGGATAATTTTTCTATATTAGGTACAATTTATCTAACTGAATTGAGAAAAAATTACAGTATTATCTCTC
25 AAAATAACATCAATCTATTTTTGTAAACCTGTTCACTATTAAATTTTGCCCTAAAAGACCTCTAATAATGATTGTTGCCAG
TGACTGATGATTAATTTTATTTTACTTAAATAAGAAAAGGAGCACTTTAATTACAACGAAAAATCAGATTGTTTGCAGTCC
TTCTTACACTAATTTGAACCTCTTAAAGATTGCTGCTTTTTTTTGACATTGTCAATAACGAAACCTAATTGTAAAACAGTCA
CATTTACTACCAATAACTTTTAGTTAATGTTTTACAAGG
- 91
- 30 MLRTAMGLRSWLAAPWGALPPRPPLLLLLLLLLLLLLLQPPPTWALSPRISLPLGSEERPFRLFEEHISNYTALLLSRDGRTLYV
GAREALFALSSNLSFLPGGEYQELLWGADAEEKKQCSFKGKDPQRDCQNYIKILLPLSGSHLFTCGTAAFSPMCTYINMENFTL
ARDEKGNVLEDGKGRCPDPNFKSTALVVDGELYGTVSSFQGNDAISRSQSLRPTKTESSLNWLQDPFAVASAYIPESLGS
LQDDDKIYFFSETGQEPFEFFENTIVSRIARICKDEGGERVLQQRWTSFLKAQLLCSRDDGFPFNVLDQVFTLSPSPQDWR
DTLFGVFTSQWHRGTTGSAVCFVTKMDVQRFVSGLYKEVNRETQWYTVTHPVPTPRPGACITNSARERKINSLSLQLPDRVL
35 NFLKDHFLMDGQVRSRMLLLQPPQARYQRAVHRVPLHHTYDVLFLGTGDRHLKAVSVGPRVHIIEELQIFSSGQPVQNLDD
THRGLLYAASHSGVVQVPMANCSLYRSCGDCLLARDPYCAWSGSSCKHVSlyQPLATRPWIQDIEGASAKDLCASSVVSFSF
VPTGEKPCQVQFPNTVNTLACPLLSNLATRLWLRNGAPVNASASCHVLPTGDLILLVGTQQLGEFQCWSLEEGFQQLVASYCP
EVVEDGVADQDEGGSVPIISTSRVSAPAGGKASWGADRSYWKEFLVMCTLFVLAVLLPVLFLLYRHRNSMKVFLKQGECAV
HPKTCPVVLPETRLNGLGPPSTPLDHRGYQSLSDSPPGARVFTSEKRPLSIQDSFVEVSPVCPRPVRVLGSEIRDSVV
- 40 92
- GCTCTGCCCCAAGCCGAGGCTGCGGGCGCGCGCGCGGAGGACTGCGGTGCCCGCGGAGGGGCTGAGTTTGCCAGGGCCCA
CTTGACCCCTGTTTCCCACTCCCGCCCCCAGGTCCGAGGCGGGGCCCCCGGGCGACTCGGGGCGGACCGCGGGCGGAG
CTGCGCCCCGTGAGTCCGGCGAGCCACTGAGCCCGAGCCGCGGACACCGTCTGCTCTCCGAATGCTGCGCACCGCG
ATGGGCGCTGAGGAGCTGGCTCGCCGCCCCATGGGGCGCGCTGCCGCTCGGCCACCGCTGCTGCTGCTCTGCTGCTGCTGCT
45 CTGCTGCAGCCGCGCTCCGACCTGGGCGCTCAGCCCCGGATCAGCCTGCCTCTGGGCTCTGAAGAGCGGCCATTCTCAGA
TTCAAGCTGAACACATCTCCAATACACAGCCCTTCTGCTGAGCAGGATGGCAGGACCTGTACGTGGGTGCTCAGAGGCC
CTCTTTGCACTCAGTAGCAACCTCAGCTTCTGCCAGGCGGGAGTACCAGGAGCTGCTTTGGGGTGCAGACGAGAGAAGAA
CAGCAGTGCAGCTTCAAGGCAAGGACCCACAGCGCAGTGTCAAACTACATCAAGATCTCTGCGCTCAGCGGCAGTCA
CTGTTACCTGTGGCAGCAGCCTTCAAGCCCATGTGTACCTACATCAACATGGAGAATTACCTTGGAAGGACGAGAAG
50 GGAATGTCTCTTGAAGATGGCAAGGGCGGTGTCCTTCGACCCGAATTTCAAGTCCACTGCCTGTGTTGATGGCGAG

CTCTACACTGGAACAGTCAGCAGCTTCCAAGGGAATGACCCGGCCATCTCGCGGAGCCAAAGCCTTCGCCCCACCAAGACCGAG
AGCTCCCTCAACTGGCTGCAAGACCCAGCTTTGTGGCTCAGCCTACATTCTGAGAGCCTGGGCAGCTTGAAGGCGATGAT
GACAAGATCTACTTTTTCTTCAGCGAGACTGGCCAGGAATTTGAGTTCTTTGAGAACACCATTTGTGTCCCGCATTGCCCGCATC
TGCAAGGGCGATGAGGGTGGAGAGCGGGTGCTACAGCAGCGCTGGACCTCCTTCCTCAAGGCCAGCTGCTGTGCTACCGGCC
5 GACGATGGCTTCCCCTTCAACGTGCTGCAGGATGTCTTACGCTGAGCCCCAGCCCCAGGACTGGCGTGACACCTTTTCTAT
GGGGTCTTCACTTCCCAGTGGCAGGGGAACACAGAAGGCTCTGCCGTCTGTGCTTTCACATGAAGGATGTGCAGAGAGTC
TTCAGCGGCTCTACAAGGAGGTGAACCGTGAGACACAGCAGTGGTACACCGTGACCCACCCGGTGCCACACCCCGGCTTGA
GCGTGATCACCAACAGTGGCCGGGAAAGGAAGATCAACTCATCCCTGCAGCTCCCAGACCGGTGCTGAACTTCTCAAGGAC
CACTTCTGATGGACGGGCAGGTCCGAAGCCGATGTGTGCTGCAGCCCCAGGCTCGCTACCAGCGGTGGCTGTACACCGC
10 GTCCCTGGCTGCACACACCTACGATGTCTCTTCTGGGCACTGGTGACGGCCGGCTCCACAAGGCAGTGAGCGTGGGCCCC
CGGTGACACATTCATTGAGGAGCTGCAGATCTTCTCATCGGACAGCCCGTGCAAGATCTGCTCTGGACACCCACAGGGGGCTG
CTGTATGCGGCTTCACTCGGGCTAGTCCAGGTGCCATGAGCCAACTGCAGCTGTACAGGAGCTGTGGGAGTGCCTCTCTC
GCCCGGAGCCCTACTGTGCTTGGAGCGGCTCCAGCTGCAAGCACGTGAGCTTACCAGCTTACCTGCCCCAGGCGCTGG
ATCCAGGACATCGAGGGAGCCAGCGCCAAAGGACCTTTGCAGCGCTCTTCGGTGTGTGCCCGCTCTTTTGTCAACAGGGGAG
15 AAGCCATGTGAGCAAGTCCAGTTCCAGCCCAACAGTGAACACTTTGGCTGCCCGCTCCTCTCCAACCTGGCGACCCGACTC
TGGCTACGAACGGGGCCCCGTCATGCTCGGCTCCTGCCAGTGTCTACCCACTGGGGACCTGTGCTGGTGGGACCCCAA
CAGCTGGGGAGTTCCAGTGTGCTCACTAGAGGAGGGCTTCCAGCAGCTGGTAGCCAGCTACTGCCAGAGGTGGTGGAGGAC
GGGGTGGCAGACCAACAGATGAGGGTGGCAGTGTACCCGTCATTATCAGCACATCGCGTGTGAGTGCACAGCTGGTGGCAAG
GCCAGCTGGGGTGCAGACAGTCTTACTGGAAGGAGTTCTGGTGTATGTGCAGCTCTTTGTGCTGGCGTGTGCTCCAGTT
20 TTATTCTTGTCTTACCGGCACCGGAACAGCATGAAAGTCTTCTGAAGCAGGGGGAATGTGCCAGCTGCACCCCAAGACCTGC
CCTGTGGTGTGCTGCCCCTGAGACCCGCCACTCAACGGCTAGGGCCCCCTAGCACCCCGCTCGATCACCGAGGTACAGTCC
CTGTGACAGCCCCCGGGTCCCGAGTCTTCACTGAGTCAAGAGAAGAGGCCACTCAGCATCAAGACAGCTTCGTGGAGGTA
TCCCCAGTGTGCCCCGGCCCCGGGTCCGCCCTGGCTCGGAGATCCGTGACTCTGTGGTGTGAGAGTGACTTCCAGAGGACGC
TGCCCTGGCTTCAAGGGCTGTGAATGCTCGGAGAGGGTCAACTGGACCTCCCTCCGCTCTGTCTTCTGTGGAACACGACCGTG
25 GTGCCCGGCCCTTGGGAGCCTTGGGGCCAGCTGGCCTGCTGCTCTCAGTCAAGTAGCGAAGCTCCTACCACCCAGACACCCAA
ACAGCCGTGGCCCCAGAGGTCTTGGCCAAATATGGGGGCTGCTTAGGTTGGTGAACAGTGTCTTATGTAACTGAGCCCT
TTGTTTAAAAACAATTTCCAAATGTGAACTAGAATGAGAGGGAAGAGATAGCATGGCATGCAGCACACAGGCTGTCTCAGTT
CATGGCTTCCAGGGGTGCTGGGGATGCATCCAAAGTGGTTGTCTGAGACAGAGTTGGAACCCCTACCAACTGGCCTCTTCACT
CTTCCACATTTATCCCGTGCACCGGCTGCCCTGTCTCACTGCAGATTCAGGACCAGCTTGGGCTGCGTGCGTTCTGCTTGGC
30 AGTCAGCCGAGGATGTAGTTGTGTGCTGGCGCTTCCAGCCTGTATCAGGCTGTGGCCACACGAGAGGACGCGGCTCAGGAG
GGTCTGGGCTCGGACCCAACTCCTGGACCTTCCAGCCTGTATCAGGCTGTGGCCACACGAGAGGACGCGGCTCAGGAG
AGATTTCTGTACATGTACGCCTTTCCCTCAGAAATCAGGGAAGAGACTGTGCGCTGCCTTCTCCGTTGTGTGCGTGAACACC
GTGTGCCCTTCCACCATATCCACCTCGCTCCATCTTTGAACTCAAAACAGGAACTAACTGCACCTGTGCTCTCCCCA
GTCCCCAGTTACCTTCCATCCCTCACCTTCTCCACTCTAAGGATATCAACACTGCCAGCACAGGGGCCCTGAATTTATGT
35 GGTTTTATACATTTTTTAATAAGATGCACCTTTATGTATTTTTTAATAAAGTCTGAAGAATTACTGTTT

92a

GGCAGAGGCGCTTGGCGGTGAGCTCTGCCCAAGCCGAGGCTGCGGGGCCGTGCTCCGGCGAACGCCACTCGAGACGGGTT
CGGCTCCGACGCCCCCGCGCGCGCGGGAGGACTGCGGTGCCCGCGGAGGGGCTGAGTTTGGCCAGCCGCGCGCCCTCCTG
ACGCCACGGGGCGCTCCCGACTCAAACGGTGGGCCCACTTGACCTGTTTCCACCTCCCGCCCCCAGGTCCGGAGGCGCC
40 CGGGTGAAC TGGACAAAGGTGGAGGGCGGGGGTCCAGGCCCTCCGCGGGGCCCGGGGCGACTCGGGGCGGACCGCGGGG
CGGAGCTGCCGCCCCCGGGGCCCCGCTGAGCCCCGCTGGCGCCCCGCTCGACGGCGGGGTGAGTCCGCGCGAGCCACC
TGAGCCCGAGCCGCGGACACCGTCCCTCACTCAGGCCGCTCGGTGAGTCCGGCTCGGCGCCCTGTGGCAGCGAGGTGCT
CTCCGAATGCTGCGCACCGCATGGGCTGAGGAGCTGGCTCGCGACGAGAGGCTTACGACCGGTGGCGCTACCCGGACTCCT
CGAECGAGCGGCCCCATGGGGCGGCTGCCGCTCGGCCACCGCTGTGCTGTCTCTGCTGGGGTACCCCGCGGACGGCGG
45 AGCCGGTGGCGACGACGAGGACGACTGTCTGCTGTCAGCCGCGCTCCGACTTGGGCTCTGAAGAGCTTGAACACATCTCAACTACACAG
AGGACGACGTCGGCGGGAGGCTGGACCCGCGAGTGGGGGCGATCAGCTGCCCTTGGGCTCTGAAGAGCTTGAACACATCTCAACTACACAG
CCCTTCTGCTGAGCAGGATGGCAGGACTTGTGTAGAGGTGTATGTGTGGGAAGACGACTCGTCCCTACCGTCCACCTGTGTA
CGTGGGTGCTCGAGAGGCCCTTTTGCACTCAGTAGCAACCTTGGGACATGACCCACGAGCTCTCCGGGAGAAACGTGAGTCA
50 TCGTTGACAGCTTCTGCCAGGCGGGGAGTACCAGGAGCTGCTTTGGGGTGCAGACGGTCAAGGACGGTCCGCCCTCATAG
TCCTCGACGAAACCCACGTCTGCCAGAGAAGAAACAGCAGTGCAGCTTCAAGGGCAAGGACCCACAGCGGACGTCTCTTCTT

TGTCGTCACGTCGAAGTTCCCGTTCTGGGTGTCGCGCTGTGTCAAACTACATCAAGATCCTCCTGCCGCTCAGCGGCAGTCA
CCTGTTACAGTTTTGATGTAGTTCTAGGAGGACGGCGAGTCGCCGTCAGTGGACAAACCTGTGGCACAGCAGCCTTCAGCCCC
ATGTGTACCTACATCAACATGGGTGGACACCGTGTGTCGGAAGTCGGGGTACACATGGATGTAGTTGTACCAGAACTTCAACC
TGGCAAGGGACGAGAAGGGGAATGTCCTCCTGGAAGATTCTTGAAGTGGGACCGTTCCCTGCTCTTCCCTTACAGGAGGACCT
5 TCTAGGCAAGGGCCGTTGTCCCTTCGACCCGAATTTCAAGTCCACTGCCCTGGTCCGTTCCCGGAACAGGGAAGCTGGGCTTA
AAGTTCAGGTGACGGGACAGGTTGATGGCGAGCTCTACACTGGAACAGTCAGCAGCTTCCAAGGGAATGCCAACTACCGCTCG
AGATGTGACCTTGTGTCAGTCGTCGAAGGTTCCCTTACACCCGGCCATCTCGCGGAGCCAAAGCCTTCGCCCCACCAAGACCGAGA
GCTGGGCCGCTAGAGCGCCTCGGTTTCGGAAGCGGGGTGGTTCTGGCTCTCGTCCCTCAACTGGCTGCAAGACCCAGCTTTGT
GGCCTCAGCCTACATTCAGGAGTTGACCGACGTTCTGGGTGCAAAACACCGGAGTCGGATGTAAGGTGAGAGCCTGGGCAGC
10 TTGCAAGCGATGATGACAAGATCTACTTTTCTACTCTCGGACCGCTCGAACGTTCCGCTACTACTGTTCTAGATGAAAAGA
TCAGCGAGACTGGCCAGGAATTTGAGTTCCTTGAGAAACACCATTTGTGTCCAGTCGCTCTGACCGGCTCTTAAACTCAAGAACT
CTTGTGGTAACACAGGCGCATTCGCCGCATCTGCAAGGGCGATGAGGGTGGAGAGCGGGTCTACAGCGTAACGGGCGTAGACG
TTCGCCCTACTCCACCTCTCGCCACGATGTGACGCGTGGACCTCCTTCTCAAGGCCAGCTGCTGTGCTACGGGCCGCG
TCGCGACCTGGAGGAAGGAGTTCCGGGTGACGACACGAGTGC CGGGCACGATGGCTTCCCTTCAACGTGCTGAGGATGTCT
15 TCACGCTGAGCCCCGCTACCGAAGGGGAAGTTGCACGACGCTCTACAGAAGTGCAGTCTCGGGAGCCCCAGGACTGGCGTGA
CACCCTTTTCTATGGGGTCTTCACTTCCCATCGGGGTCTGACCGCACTGTGGGAAAAGATACCCAGAAAGTGAAGGGTGTGG
CACAGGGAACTACAGAAGGCTCTGCCGTCTGTGCTTCAATGACACCGTGTCCCTTGATGTCTTCCGAGACGGCAGACAC
AGAAGTGTACTAGGATGTGACAGAGCTTTCAGCGGCTCTACAAGGAGTGAACCGTGAGTCTTACAGTCTCTCAGAAGTC
GCCGGAGATGTTCCTCACTTGGCACTCACACAGCAGTGGTACACCGTGACCCACCGGTGCCACACCCGGCTGGTGTGTC
20 GTCACCATGTGGCACTGGGTGGGCCACGGGTGTGGGGCCGACAGCGTGATCAACACAGTGC CGGGAAAGGAAGATCAAC
TCATCCTGTCTGCACGATAGTGGTTGTACGGGCCCTTTCTTCTAGTTGAGTAGGGACGAGCTCCAGACCGCGTGCTGAAC
TCTCAAGGACCATTCTGTGAGTCTGAGGGTCTGGCGCACGACTTGAAGGAGTTCTTGGTGAAGGACTACCTGGGGCAGGT
CCGAAGCCGCATGTCTGTGTCAGCCCCAGGCTCGCTACCAACCGTCCAGGCTTCGGCGTACGACGACGAGCTCGGGTCCGA
GCGATGGTGGCGTGGCTGTACACCGCTCCCTGGCTGCACACACCTACGATGTCCCGCGCACCGACATGTGGCGCAGGGAC
25 CGGACGTGGTGTGGATGCTACAGGCTTTCCTGGGCACTGGTGACGGCCGGCTCCACAAGGCAGTGAGCGTGGGCAGAAGGACCC
GTGACCACTGCCGGCCGAGGTGTTCCGTCACTCGCACCCGCCCCGGGTGCACATCATTGAGGAGCTGCAGATCTTCTCATCGGG
ACAGCCGGGGCCACGTGTAGTAACCTCTCGACGCTCTAGAAGAGTAGCCCTGTTCGGCGTGCAGAATCTGCTCTGGACACCCAC
AGGGGGCTGCTGTATGCGGCCCTGCACGCTCTAGACGAGGACCTGTGGGTGTTCCTCCGACGACATACGCCGACACACTCGGGCG
TAGTCCAGGTGCCATGGCCAACTGCAGCCTGTACCGGGTGTAGCGCCGATCAGGTCCACGGGTACCGGTTGACGTCGGACAT
30 GGCCAGCTGTGGGACTGCCCTCTCGCCGGGACCCCTACTGTGCTTGGAGCGGTGCACACCCCTGACGGAGGAGCGGGCCCTG
GGGATGACACGAACCTCGCCCTCCAGCTGTAAGCACGTCAGCCTCTACAGCCTCAGCTGGCCACCAGGCGAGGTGCAGATTCTG
TGCAGTCGGAGATGGTGGAGTCGACCGGTGGTCCGCTGGATCCAGGACATCGAGGGAGCCAGCGCAAGGACCTTTCGACGCG
CGGCACCTAGGTCTGTAGCTCCCTCGGTGCGGTTCCTGGAAACGTCGCGCTCTTCGGTTGTGTCCCGCTCTTTGTACCAAC
AGGGGAGAAGCCATGTGAAGAAGCCAAACACAGGGGCAGAAAACATGGTTGTCCCTCTTCGGTACACTGCAAGTCCAGTTCAG
35 CCCAACACAGTGAACACTTTGGCCTGCCCGCTCCCGTTCAAGTCAAGGTGGGTGTGTACTTGTGAAACCGGACGGGCGAGG
TCTCCAACTGGCGACCCGACTCTGGCTACGCAACGGGGCCCCGTCATAGAGGTGGACCGTGGGCTGAGACCGATGCGTT
GCCCCGGGGCAGTTAGCCTCGGCCCTCTGCCACGTCTACCCACTGGGGACCTGCTGCTGGTGGGCGGAGCCGGAGGACGGTG
CACGATGGGTGACCCCTGGACGACGACACCCACCCCAACAGCTGGGGGAGTTCCAGTGCTGGTCACTAGAGGAGGGCTTCCGT
GGGTGTGACACCCCTCAAGGTACAGACAGTATCTCTCCGAAGGAGCAGCTGGTAGCCAGTACTGCCAGAGGTGGTGG
40 AGGACGGGGTGGCATCGTCGACCATCGGTGATGACGGGTCTCCACCACCTCTGCCCCACCGTGACCAACAGATGAGGGTGG
CAGTGTACCCGTCAATTATCAGCACATCGCGCTGGTTGTCTACTCCACCGTCACATGGGCGAGTAATAGTCGTGTAGCGCTGTG
AGTGCACACAGCTGGTGGCAAGGCCAGTGGGGTGCAGACAGGTCTTACACTCAGTGGTTCGACACCGTTCCGGTCGACCCAC
GTCTGTCCAGGAACGGAAGGAGTTCTTGGTGTGTCACGCTCTTTGTGCTGGCCGTGCTGTGACCTTCTCAAGGACCACTA
CAGGTGCGAGAAACACGACCGGCACGACCTCCAGTTTATTTCTTGTCTTACCGGCACCGGAACAGCATGAAAGTCTTGAGGGT
45 CAAAATAAGAACGAGATGGCCGTGGCCTTGTCTACTTTTCAAGACCTGAAGCAGGGGAATGTGCCAGCGTGACCCCAAGACC
TGCCCTGTGGGACTTCTGTCCTTACACGGTTCGACGTGGGGTCTTGACGGGACACCTGCTGCCCTTGAGACCCGCCAC
TCAACGGCTTAGGGCCCCCTAGCACACGACGGGGGACTCTGGGCGGGTAGTTGGCGGATCCCGGGGATCGTGCCACTCGA
TCACCGAGGGTACCAGTCCCTGTGACAGACCCCCGGGGCGGTGAGCTAGTGGCTCCCATGGTACGGGACAGTCTGTCCGGG
GGCCCCCGCGAGTCTTCACTGAGTCAGAGAAGAGGCCACTCAGCATCCAAGACAGCTGGCTCAGAAGTACTCAGTCTCTTCT
50 CCGGTGAGTCTGATGTTCTGTGATCGTGGAGGTATCCCAAGTGTGCCCCGGCCCCGGGTCCGCTTGGCTGAGACCTCCA
TAGGGGTACACGGGGGCCGGGGCCAGGCGGAACCGAGGAGATCCGTGACTCTGTGGTGTGAGAGTACTTCCAGAGGACG
CTGCCCTCTAGGCACAGAGACACCACTCTGACTGAAGGTCTCTGCGACGGGTGGCTTCAGGGGTGTGAATGCTCGGAG

- AGGGTCAACTGGACCTCCCTCACCAGTCCCGACACTTACGAGCCTCTCCAGTTGACCTGGAGGGGAGCGCTCTGCTCTT
CGTGGAACACGACCGTGGTGGCCCGCCCTTGGGAGCCTGCGAGACGAGAAGCACCTTGTGCTGCGACACCGGCCGGGAACCTT
CGGATGGAGCCAGCTGGCTGCTGCTCTCCAGTCAAGTAGCGAAGCTCTACCAACCTCGGTGCGACCGGACGACGAGAGGTGAG
TTCATCGCTTCGAGGATGGTCCCAGACACCCAAACAGCCGTGGCCCCAGAGGTCTGGCCAAATATGGGGGGTCTGTGGGTTT
5 GTCCGACCCGGGTCTCCAGGACCGGTTTATACCCCGCTGCTAGGTTGGTGGAACTGCTCCTTATGTAACCTGAGCCCTT
TGCGGACGGATCCAACACCTTGTACAGGAATACATTTGACTCGGGAACTTTAGAAAACATTCCAAATGTGAACTAGAA
TGAGAGGAAGAGATAGCAAACTTTTGTAAAGTTTACATTTGATCTTACTCTCCCTTCTCTATCGATGGCATGCAGCACAC
ACGGCTGCTCCAGTTTATGGCTCCAGGGGTGCTACCGTACGTCGTGTGCGGACGAGGTCAAGTACCGGAGGGTCCCCACG
TGGGGATGCATCCAAAGTGGTTGTCTGAGACAGAGTTGGAAACCTCACCACCCCTACGTAGGTTTACCAACAGACTCTGTCT
10 CAACCTTTGGGAGTGGAAGTGGCTCTTACCTTCCACATTATCCCGCTGCCACCGGTGCCCTGTTTGACCGGAGAGTGGAA
GGTGAATAGGGGACGGTGGCCGACGGGACACTTCTGACAGATTACAGACAGCTTGGGCTGCGTGCCTTCTGCTTGGCAGA
GTGACGCTAAGTCTGGTTCGAACCCGACGACGCAAGACGGAACGGTGTGACCGAGGATGTAGTTGTGCTGCGGTGCTCCC
ACCACCTCAGGGACAGTCCGGCTCTACATCAACAACGACGCGACAGGGTGGTGGAGTCCCTGCAGAGGGCTAGGTTGGCACT
GCGGCCCCACAGGTCTTGGGCTCGGACCGTCTCCCGATCCAAACCGTGACGCGGGAGTGGTCCAGGACCGGAGCTGGCAAC
15 TCCTGGACCTTTCCAGCCTGTATCAGGCTGTGGCCACACGAGAGGAGTTGAGGACCTGGAAAGGTTCGGACATAGTCCGACACCG
GTGTGCTCTCCTCAGCGGAGCTCAGGAGAGATTTCGTGACAATGTACGCCTTTCCCTCAGAGTCGCGCTCGAGTCTCTCTAA
AGCACTGTTACATGCGGAAAGGGAGTCTATTCAGGGAAGAGACTGTGCGCTGCCTTCTCCGTGTTGCGTGAGAACCTAAGTC
CCTTCTCTGACAGCGGACGGAAGGAGGCAACAACGCACTCTTGGCGTGTGCCCCCTTCCACCATATCCACCTCGCTCCATCTT
TGAACCTCAAAGCACACGGGAAGGGTGGTATAGGTGGGAGCGAGGTAGAACTTGAGTTTACGAGGAACCTAAGTGCACCTGG
20 TCCTCTCCCCAGTCCCACTTACCCGTCTCTTATTGATTGACGTGGGACAGGAGAGGGGTACGGGTCAAGTGGGTCCATCCC
TCACCTTCTCCACTCTAAGGATATCAACACTGCCCAGCACAGGTAGGAGTGAAGGAGGTGAGATTCCCTATAGTTGTGAC
GGTTCGTGAGGGGCCCTGAATTTATGTGTTTATATATTTTAAATAAGATGCACCTCCCGGGACTTAAATACACCAAAAA
TATATAAAAAATATTCTACGTGATTATGTCATTTTAAATAAGTCTGAAGAATTACTGTTTAAAAAATAATACAGTAA
AAAATTATTTCAGACTTCTTAATGACAAATTTTATTTTAAAAAAACTTTTGTG
- 25 93
MVGYDPKPDGRNNTKFQVAVAGSVSLVTRALISPFVIRFQQLQHERLSRSDPSAKYHILQASRQILQEEGPTAFWKGHVP
AQILSIGYGAQVPLSFEMLTSLVHRGSVYDAREFSVHFVCGGLAACMATLTVHPVDVLRTRFAAQGEKPVYNTLRHAVGTMYRS
EGPQVYFKGLAPTLIAIFPYAGLQFSCYSSLKHLKWAIPAEGKKNNENLQNLGSGAGVISKTLTYPLDLFKRLQVGGFEHA
RAAFGQVRRYKGLMDCAKQVLQKEGALGFPGKLSPLSLKALSTGFMFFSYEFFCNVFCMNRRTASQR
- 30 94
ATGGTTGGCTATGACCCCAAACAGATGGCAGGAATAACACCAAGTTCCAGGTGGCAGTGGCTGGGTCTGTGCTGGACTTGT
ACTCGGGCGCTGATCAGTCCCTTCGACGTCATCAAGATCCGTTTCCAGCTTACGATGAGCGCCTGTCTCGCAGTGACCCAGC
GCAAAGTACCATGGCATCTCCAGGCTCTAGGAGATTCTGCAGGAGGAGGGTCCGACAGCTTTCTGGAAAGGACAGTCCCA
GCTCAGATTCTCTCCATAGGCTATGGAGCTGTCAATTTCTGTCATTTGAAATGCTGACGGAGCTGGTCCACAGAGGACGCTG
35 TACGACGCCCGGAATTCTCAGTGCATTTGTATGTGGTGGCTGGCTGCCTGTATGGCCACCTCAGTGTGACCCCGTGGAT
GTTCTGCGCACCCGCTTTGCAGCTCAGGGTGAGCCCAAGGTCTATAATACGCTCGCCACGCGCTGGGGACCATGTATAGGAGC
GAAGGCCCCAGGTTTTCTACAAAGGCTTGGCTCCACCTTGATCGCCATCTTCCCTACGCGGGCTCGAGTTCTCTGTCTAC
AGCTCCTTGAAGCACTGTACAAGTGGGCCATACAGCCGAAGGAAAGAAAAATGAGAACCTCCAAACCTGCTTTGTGGCAGT
GGAGCTGGTGTATCAGCAAGACCTGACATATCCGCTGGACCTCTTCAAGAAGCGGCTACAGGTGGAGGGTTTGGAGCATGCC
40 AGAGCTGCCTTTGGCCAGGTACGGAGATACAAGGCCCTCATGGACTGTGCCAAGCAGGTGTACAAAAGGAAGGCGCCTGGGC
TTCTTCAAGGCCGTGTCCTCCAGCTTGTGAAGGCTGCCCTCTCCACAGGCTTCATGTTCTTCTCGTATGAATTTCTTGTAA
GTCTTCACTGCATGAACAGGACAGCCAGCCAGCGCTGA
- 95
MELALRRSPVPRWLLLLPLLLGLNAGAVIDWPTEEGKEVWDYVTVRKDAYMFWWLYATNSCKNPFSELPLVMWLQGGPGSSSTG
45 FGNFEEIGPLSDSLKPRKTTWLQAASLLFVDNPVGTGFSYVNGSGAYAKDLAMVASDMMVLLKTFPSCHKEFTQVPFYIFSES
GGKMAAGIGLELYKAIQRTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVLNAVNGLYREATL
WGKAEMIIBQNTDGVNFYNTKSTPTSTMESSLEFTQSHLVCLQRHVRHLQRDALSQLMNGPIRKKLKIIPEDQSWGQATN
VFVNMEEDFMKPVISIVDELLEAGINVTYNGQLDLIVDTMGQEAWVRKLKWPBPKFSQLKWKALYSDPKSLETSAFVKSYKN
LAFYWLKAGHMPVPSDQGMALKMMRLVTQOE

96

CCTGTTGCTGATGCTGCCGTGCGGTACTTGTTCATGGAGCTGGCACTGCCGGCGCTCTCCCGTCCCGCGGTGGTTGCTGCTGCTGC
CGCTGCTGCTGGGCCTGAACGCAGGAGCTGTCAATTGACTGGCCACAGAGGAGGGCAAGGAAGTATGGGATTATGTGACGGTCC
GCAAGGATGCCCTACATGTTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCTGGTCAATGTGGC
5 TTCAGGGCGGTCCAGGCGGTCTAGCACTGGATTGAGAACTTTGAGGAAATTTGGGCCCTTGACAGTATCTCAAACCACGGA
AAACCACCTGGCTCCAGGCTGCCAGTCTCCTATTTGTGGATAATCCCGTGGGCACTGGGTTCAGTTATGTGAATGGTAGTGGTG
CCTATGCCAAGGACCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGGCACAAGAAATTCAGAG
CAGTTCCTATTCTACATTTTCTCAGAGTCTATGGAGGAAAAATGGCAGCTGGCATTTGGTCTAGAGCTTTATAAGGCCATTTCAGC
GAGGGACCATCAAGTCAACTTTGCGGGGGTTGCCTTGGGTGATTCCCTGGATCTCCCCCGTGGATTTCGGTGTCTCTCTGGGGAC
10 CTTACCTGTACAGCATGTCTCTTCTCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTACTGAATGCCGTAA
ATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAATGATCATTGAACAGAACACAGATGGGGTGAATCTCT
ATAACATCTTAACATAAGCACTCCACGTCTACAATGGAGTCGAGTCTAGAAATTCACACAGAGCCACCTAGTTTGTCTTTGTCT
AGCGCCACGTGAGACACCTACAACGAGATGCCTTAAGCCAGCTCATGAATGGCCCCATCAGAAAGAAGCTCAAAATTTATCTCTG
AGGATCAATCTTGGGAGGCCAGGCTACCAACGTCTTTGTGAACATGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTTGTGG
15 ACGAGTTGCTGGAGCCAGGATCAACGTGACGGTGTATAATGGACAGCTGGATCTCATCGTAGATACCTAGGGTCAGGAGGCCCT
GGGTGCGGAACTGAAGTGGCCAGAACTGCCCTAAATTCAGTCAGCTGAAGTGGAGGCCCTGTACAGTGACCCATAATCTTTGG
AAACATCTGCTTTTGTCAAGTCCCTACAAGAACCTTGCTTTCTACTGGATTCTGAAAGCTGGTCATATGTTCTCTTGACCAAG
GGGACATGGCTCTGAAGATGATGAGACTGGTGACTCAGCAAGAATAGGATGGATGGGGCTGGAGATGAGCTGGTTTGGCCTTGG
GGCAGAGAGCTGAGCTGAGGCCGCTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCTAACTGGGGCTGTGATCAAGAGGTT
20 CTGACCAGCTTCTGAGAGGATAAAATCATTTGTCTCTGGAGGCAATTTGGAAATTTATTTCTGCTTCTTAAAAAAACCTAAGATT
TTTTAAAAAATTGATTGTTTGTGATCAAAATAAAGGATGATAATAGATATTTTCTTATGACAGAGCAAAATGATGTGAT
TTATAGAAAACTGGGAAATACAGGTACCCAAAGAGTAAATCAACATCTGTATACCCCCCTCCAGGGGTAAGCACTGTTACCA
ATTTAGCATATGTCCTTGACAGAAATTTTTTTTCTATATATACATATATATTTTTTACCAAAATGAATCATTACTCTATGTTGTT
TTACTATTTGTTTGACATATCAGTATATCTGAAACACCTTTTCATGTCAATAAATGTTCTTCTTAACATTAA

25 98

AATAGTTTCATTAATTAGTCTTTATTATGTGAAAAGACTTTTCTAAAAATTTCTCCATAATAATTGAAGATATAAATTACAATAAT
TAAAAATTTACAAAAGAATAAAAGCATTAAAAATAAGTGTGAAAAATTTAAAAAGGGTAAAGATGAAAATTTATTTAATTTTAAT
CTCAAAAATCCGAAAGTGATTTTATCAATGTCTTAAGAAAATTTAAAAAGAAAACAACTTAATAATGTAAAGAAAGGAGGTTTC
TGCCGGGCTGGGGGCGGGAGCTACGGCTTCTCTGGGGACGCGGAAGCGAGAAGCAGGGACCTTGGGGCGCGCCCGGTTTTCGG
30 GACGGTGCCCCGGCCCCCAGCCACGCCCTCCGGCGCTGGCATGGTGTGCTTGTCTGGGACCCGGCCGAGGGTGGCTGGGAGC
GCTGCATGACCCCGCCACGCGCTCCCACTCCTAGGCGCACAGGTCGAGGAGGACCGCGCTGACTACAAGAGTTCCAGGGCC
CCACACCTGCAGCATAGCTGGGACGACTCTGTACCTTTTCCAGGACGAGGAGGACGCGCGTCGAGGGCTTGGAGAGCGTCC
CGGAGGAGGGCGTCCAGGAGGCATGGGGTGTGCGGACGCTGCTGCGGCGAGGGGTGAGCGTTGAGAAGGCGCAGGAGACTGA
CCACAATGGCTGGACACGCTTAATGTCCCTGTACCAAGGCTTTGTGGATATACCGTGGTGGCTTAGCTGAGTAACCCAC
35 ACTGACGTCAACTGGCAGGACAGGGAGGGGAACGAGCCCTAAATACAGCTGCACAGGACGAGCTGCCTCTGGCCCCATGGGC
ACCGTTGCTTGGACACTGGCTCTCAGCTTGTGCTTCTCTGGTCAGCATCACGTGGGCTGGTTTCAATTTAGCTTTTAACACAGAT
CTGGTGGAAACACAGGCTGCCAGGCCCTGCCCTGGAGTCTTTCTTTGAATAGGCTGGGGTGGGGCCCAAGAATGAGCAGAAG
AACAGGTTTCTTGGTGAGGCTGATGCCGCTGGCCAGGGTTCTCACGTTGAGGACCGAGAGCTTTGAGGTTTTCATACCTGATG
ATGTCCAGGAACCTTCTCAGTAGGCACTGAAGACCATCAGCAGAATCACAGACCCAGGAGAGATGTCGTGAGACAGACAG
40 AGGCATCACCGAATTAAGTGAATGAAGAAAGGAGCTGAGCATCTGTTTCAAGACTTTCGTGGCTGTTTACTAAAGAGGCT
ATCCTGGCCAGTCAGGCGACGCTATCATCACTAATCTGTTGAACCTACGTTTTCAGGCTTTCAGCTTGAAGGAGGACGCGT
TCGGGTTGAAAGCCGCCATGCAGGCTCGAACGGATCGCATCGGAGCCCTGAGCTAGCAGGGGCGGGTGTCCACGCGAGGGCCCC
CGCGGTGGGATGTGCGCAGAGGAGTGGCCACTTATACGTCCGCTCATGAGAGGCTCCTGGAGCGCCCCCTGCCGGGAGCTGTT
GGGGAAGAAAGTACCAGCTTGAAGCTGCCGCGCTCCCCGAGAGGCGCAAAAGCCGAGGGCTCCAAGAACTGCCTGCAGAGGCT
45 CAGGAGACTGCGAGCTGTCCGCGCTGACGCGCGCTCCATGCGGGGCCAGAGGACCGGGCGCCCTGGACCACATGGTCAGGAT
GACCACGAGCTCTAACAGCCCCCGGGGGCCCTCGCGTGCCAGACTATGTGCCCCGAGAGCACCTTGTGCTGG

99

MAPMGIRLSPLGVAVFCLLGLGVLYHLYSGFLAGRFSFLGLGGEPPGGGAAGPAAAADGGTVDLREMLAVSVLAAVRGGDEVRRV
RESNVLHEKSKGKTREGAEDKMTSGDVLNRMKPYLLKTAFPSPQINTEHVDAADQEVILWDHKKIPEDILKEVITPKVEPAES
50 VTVWIDPLDATQEYTEDLRKYVTMVCVAVNGKPMGLVHKKPFSEYTAWAMVDGGSNVKARSSYNEKTPRIVVSRSHSGMVKQV

[illegible]

25 TTTT TTTT TTTT ATTT GGAT TCT ATTT ATTT TAAA ATTT AAAT TCT TTT TGGCAA ATTT GAAC TTT CTCA ACTAAA CGTG TCCAT GTG CAGAA
TTT TAACT GTT TAGC AGGT AGT TTT GTGGCAA AGAT GGCT AAATA ATGA AGCAA ATTT AGA ATCT GTGT GTAT ACTA ATG AGCT GTCT
TTT TTT CTGT TGAG ACTAT CATT ATTT TGCT TATT ACCCA AGAGG CAATTT ACCT GAATTT GGAT GTCT GAATTT ATA ACTT ATGC
AGGA ATAG TTCT GTAAA ATAC ATTT AAAATA AACT GTAA AGAT ATTTT AATAA ATATAG TATTT ATACTA ATCT GTGT GTCTTCT TTTT
GGTTT GAATAG TAACT AAATGAG ACAC CAGCCC TTGAC ATTTG AGTTT GTTTGGT CACTAT CAGGT CCTCAT TTTCCA AGCCT CCTA
30 GTCAT TCTAG CACTG ATTTAT ATGCT GCTACT TTTAA CTGGCT CCAGCT GCCTT CACTAC ATCAG TTTAG CTTCCT CAGAAATTCAT
CAAAA TGGAC GCGCA ATTTAA ATGTAA ATTTAT AGAACT TTTTCC CAGCTG AGGCTTT TGCAC CTTCCGT ATAGT ATAGAG GGAAGC
TACAAAAA

35 MPIVD FHLQKL PAAAAAPFLDKSHSSSPGSPKTAELFFTTERTPWFHQFPQIPVAPDLSSRGTRRKDFLSTRRLGAAKVPVVSF
YHSRGVLLKASERASGGSKAVOPGRKEKETPRHRKPRFSRRKNR

AATATTGCAGAGCAACTTGAAAGGGGGAGCTGGCCGCGTAGGAGGTGAGAGCCAGCCGCAAAAAATCCGTTCCGGCGGCATTTTCT
 CTTCAAACCCGCCCGGTGCGATGTGGCATTTGAAGAAAGAAGCGCTGGGTCAATAATCAACCCGCACTTTCTCTCCAACTGC
 TGACGCGACTCTCACCGCCTTCTCAGCTAAGCCAGCTGCCGGGAAGGCCAGCTCTCCGTAAGAACCTGCCAGGGGAAGGGATTC
 40 TCTGGGCCAAGGGGACCGAGGGCTCCCGAGCAAGGGCTGGGGCGCCCTGAAAGCATACCTTTCCACTGGCGCCAAGACCCTGCC
 GGGCAAAACGCGGGGACAGACCTGAGGAGCCGTGGGAGCCCAGGGAGATCTGGGGCGGGAGCTTCGTTCCGGCCAGGGCCCTCT
 ATCCGCCCCGATACGTGGCTCTCTGGGCCGCGGGCCTCAGCGGGCAGAAGTGGGGCTCCGACCTCAAGATGAAGGTCCAGCGTCC
 TGGCATGCAGATGGTTCGGTGCCCCAGCCTCAAGCCGCTCTGCTTAAATCCAAAGGACTGGCTTTGGAGCAAGCGTAGGCGAGAT
 TATTTTAAATAATAATTTATGCTGGCGGAAGGTTTTTATTTGTTGGTTTGCTATTGGTGTTTTGTTTTTTAAGCGAAAAATGCC
 45 GATTGTTGATTTCCACCTCCAGAAATTACAGCGCCGCCGCTGCGCCCTTTCTAGACAAATCCCAAGCAGATTCACCCGCTGC
 CCCCCAAACCGCAGACTTTCTTACTGAGAGGACCCCTTGGTTCCACCAATTCCTCCCAATCCCGCTGGCTCCGATCTCTC
 ATCCAGGGGGAGTAGAAGAAAGATTTCTTCAGCACTAGAAAGATACTTGGAGCAGCAAGAGTGCCCGTGGTCTTTTCTACCA
 TTCAAGGGGTGTCTCTCTGAAGGCGCTCTGAGAGGGCCTCTGGCGGTTTCAAGGCTGTGTCAGCGGGGAGGAAGGAAAAAGGAAA

- GCCAAGACATAGAAAACACGCTTTTCCCGTAGGAAGAACCAGATGATGAGCCCTGATGAAAAGAAGGAAGACCCCGCTGCTCG
CGAAGGCCATAAAGGTTAGAAATACAGCAATTTTCTCATGCCGAAGGAAAGGCAAAATGGTCTAAGAACTGCAAGAGTGTGAT
ACAAGTTTTCCTCAACAGCGAGCTGGGACTGGCGACAATTTCTTAACTCACAGCCTCCGCAGTCGCCTTAGTCGACAAAAGAGAG
CAGCCCACGTAGCGGCCCGCGCCCGTGGCCCTTACTTGCTTTTGCCCTGAAGTTCAGTGTAGTGGTTGTTTGGGG
5 CTTGCAGGAAGAAACAAACACACTCCCTCTGACTCTCCACAACCCATCTTTGCCCTCTGCAACCGCTTCTCTGCCCATCGCTG
GTTTTGGGGTAGGCTCTTTGTCTGCCTGAGGCCCGCACTCGCAGGGCAAGAAGCTTAAAGGAATCCAACCTCCAACCTCCTCCG
AAGTTAAATATTGACCAGAACCCTTTTCAACAGGCCGCGCTTCCAGGAACCGTGGAGGGCCAACTCCTCCCAACCGCCCTGGT
GCAAAGTCCCACGCGCGAAGAGTTTGGAGCAGCGCTTACCTAGAAAGATGTTAAATCTGAACCAGGAATTGTCTCCAAC
CCAGGCGCTCAGGGAATCGCCTTTCCGGTGTCCAGGCCCTCTGCAGACAAAATAAACAGCAGAAGGTGAATGGCGGCTGCGAGG
10 GAAGGGAGTTGGAGGCCGGAGTGGGAAGCGAGGCTCATGCGGGCGCCAAAGGTGGCCCTGGGGGTGACGCGAATCAGGGAGT
CACCAGGCACTCGTCCGTGTGGCGCACTCATTTGCTCGAAATGTTCAGTCTACAAATGGTCCGAAGATGGTGATTTATTTTGT
TAAATAAATACATACTCCATCTCCACACC
- 106
- TTTTTTTTCCTAGTTTCATTAAAGATGGAGTTCATATTTGTTTCTCATATTAATTTTAAAGACAAAAAAGCAGAAGTGT
15 CACCCATTAAGTTAAAACTGGAATTTTACCTTTCCAACTTAATTTCTACCTTGCCCTTTTCATGTTAATACCATGAAGTAT
TCAGTAATCTCTTCAAGCTCCATAATTTCTACTTTCTACTTTTGCTCATTATTTCCATTCCAAGGGAAGTCTGTGTGCGT
GTCTATTTCTATGTTCAAATTACACCTGTAAGATCCAACCTCACAGGGCCATTGCCCTCTAGGGGTCTCTCTGGATCCACTTCA
ACAGGAGGTAATTTTATCTNCACCGATACCATTTCTCTCTGGACACTTACCTCTTTTANGGGGACTCCCAGGGTTGAGGG
GNG
- 20 108
- GCTTACCTGGAAAAGTGAACAACTCTCCAGGTTGACCTTTCTGTCTCCCTCTAAACACATCACATTGGCTAAAAATATCTTTTG
TTAATAAGGATTATATTACATATAAAAGTATTTAAATATTTTAAAGTACTTTACCTTAGGATAGGGTCTGAACCTCTTAGCAG
GCCTTCATGATATATTAACCTTGTGACCTCTTTGGTCTCTCTCAGTACTTCTCAATTCATCTTTGCTCTTCAGCCACAGTAAG
TTAGTTTTCCTCCCTGAGTATGTTATCAGCTCTCTGGATTCTGGGCTTTGCTCATATTTCTCTATTTCAAATACTCTTATAC
25 TTCATGTCCTTTCTTGCCACACCCCAACCCCTTCCACAGCATATGTTATGTATGATACACAATTACTCATCCTTCAGATCAGATT
CAGTTTAAATGGCATTTCTTTGGGAAGCATTTCTGATTTCTCTACCTCCAGGCTGGGTTGATTGCATGTATAGCACCTTCTAC
TGATGATACCTTTTCATGGCATTTACTTACCATAGCCTATTTGTTATTGCCCTCAAGCTCCATTAGGGCAAATTTGTTTTTCAT
TGTTTTAAAGTCAGCACCTAAATCTCCCTGGCC
- 110
- TCCTACAAATNNTATAACAAAACCTTTTATAGGAAAGTACCACAAATATGTTAAACAGAAGAAGCAAATCTCAAATAATGTGCA
30 AAAGCTGTGAGTTTCTTCTTACATAAACTGGTACCTAAGCAAGTGAGGTTTCATTTTATTTTCACTACCAATCCCCATATC
ATTATACAGTAACACCATACAGCCAAAACGGCCATGATATTCCTCCCTTCTCAGCCAAAATTGGGCAAGAGAGATGACCTTG
TAGGGGAAAAGAAACCTCTACGATAAACTGAAATGCCACCATCAGGGTTTGTGAAACTGTAGGNACAGGGTCTACGNCTCACT
TAGCTGCTAATGAGTTTCTATGGTTCCAGNTGGGGTAGTTCAAAGTA
- 35 112
- TAAAAAGCATTAGGCATATAAATGTATAAATATATTTTATCATGTACAGTACAAAAATGGAACCTTATGCATGGGCCTTAGGAA
TACAGGCTAGTATTTTCAGCACAGACTTCCCTGCTTGAGTTCTTGCTGATGCTTGCAACCGTGACAGTGGGCACCAACACAGACGT
GCCACCAACCCCTGCACACACCACCGCCACCAGGGGCCCTTGTGCGCTTGGCTTTATAAATCTCTCGGGGTGATATT
GGTGGTGATCAGCTCCTAGCATAAATGAGAGTTCCATTGGTATTGTACACGTCTCTGCTCGCTTGGGTTGCCATGTTTG
40 AGCGATGGCCCTGTGATTTACCCCTGCCTTTTACTGAATCTGTAAATTTGTTGTGCAATTGTGGTTATAGTAGACTGTAGCACA
TTGCCCTTTTCTAACTGTACATGTTTATAATCTTCATTTTAAAGTATGTGTAATTTTAAAGTATGTATTCTATTATATG
GTCTGCTTGTGAGTGGCCAGACTTGTCTACTATATTCCTTTATAAATAGCTAGCCACTTCTGGATTCTTTAGTAATGTGCT
GCATGCAAGAACTTTCCAGTAGCAGTGAAGGAGGCTGCCTTCCAAGCTTCTAAGGGATGCTGCCCTGTGTGGGGATGCATT
GCAGAGGCACTAGTAGCATGGGGCTAGAGTGGGGAGCGAGATGTAAGGGTGGGGGATAGGAGAATTCAGAGTGTCTCCA
45 GCATTAGGGTCTGAGAACTTCTGAGTTTCAAGAAACATGCAAGTGACTAACAAAATAGCTACTTACCTTTGCAGTTCTACAG
ACCTTGGGAGCTGCTTTGGGAGTGAGAAAGCAACCTTCAATGTGTTTCAACTTTAAATGTTGAATTTTTCAGACATGGT
ATCTCATTTATTTCTCTTTCTAGCGTTTGTGTAATTTCAAGCAGAATGTCTTACAGACTGTCTAGAACAGATATCATTTTA
ATCTGAAACAGCTGAGGAAGGGACAGAGAAGGTACAAGGCAAGGCAGCACAACAGATCAGGAGAATGAAGAGGGAATGCTT

118

TTCTATCAGAAATGTATTTATTTACCAGCATTTATGGCAAATGGCCAGACCTGGGCAGGGTGGTAGCAATGATAGATTTTACA
AATTAAGGCATTTCTTTTATTTCAAANTTTTCATTTACCNATATTTTCCTTGGTCTTAGAAATATGAATANTTTAGAGCTGGC
AATACTCACCATCAGGATATAATAAACGGAGGTTTCTTTNTCCGAAATCCATAAANTGTTAGTANTACTCTNTTGTACTTTTAAA
5 ATTCNATTTTTCAGTNGGCTTCCTCTCAGTGATTTAGTTAGGNAGTTTGGTACATTTTGGNGGGGTCATAACATGTTTCAT
AGGATAGGAGTACNGGGCATTNTTGGAAAAACCCG

119

MERVKMINVQRLLEAAEFLERRERECEHGYASSFFSPMSPRLQHSKPPRRLSRAQKHSSGTSNTSTANRSTHNELEKNRRAHLR
LCLERLKVLIPLGPDCTRHITLGLLNKAKAHIKKLEEAERKSQHOLENLEREQRLKWRLEQLQGPQEMERIRMSIGSTISSD
10 RSDSEREELVDVESTEFSHGEVDNISTTSISDIDHSSLSIGSDEGYSSASVKLSFTS

120

AGATTATGATCGCTGAGGCCCTCTCTACCCAGATACCGATGTTATACTGATGTGTTTTCTTTTTTTTTTTTTTTTTTTA
AGTAATTAAGGGTAGTTAAATTATTTAAAGTATACAAAGTCCAAACAGCCAGGGTAAGGTCTCCAAGAGGCCTTCCAGGGTA
AGGGAGTGGGAGAGGCCCGGTGCGCACCCGCGGTGCCATGGAGCGGGTGAAGATGATCAACGTGCAGCGTCTGCTGGAGGC
15 TGCCGAGTTTTGGAGCGCCGGGAGCGAGAGTGTGAACATGGCTACGCCCTTTCATTCCCGTCCATGCCGAGCCCCGACTGCA
GCATTCAAAGCCCCACGGAGGTTGAGCCGGGCACAGAAACACAGCAGCGGGACGAGCAACACCAGCACTGCCAACAGATCTAC
ACACAATGAGCTGGAAGAATCGACGAGCTCATCTGCGCTTTGTTTGAACGCTTAAAGTTCTGATCCACTAGGACCAGA
CTGCACGCGGCACACAACACTTGGTTTGTCTCAACAAAGCCAAAGCACACATCAAGAACTTGAAGAAGCTGAAAGAAAAAGCCA
GCACCAGCTCGAGAATTTGGAACGAGAACAGAGATTTTAAAGTGGCGACTGGAACAGCTGCAGGGTCTCAGGAGATGGAACG
20 AATACGAATGGACAGCATTTGGATCAACTATTTCTTCAGATCGTTCTGATTCAGAGCGAGAGGAGATTGAAGTGGATGTTGAAAG
CACAGAGTTCTCCCATGGAGAAGTGGACAATATAAGTACCACCAGCATCAGTGACATTGATGACCACAGCAGCCTGCCGAGTAT
TGGGAGTGACGAGGGTTACTCCAGTGCCAGTGTCAAACTTTCATTCACTTCATAGAACCAGCATGACATAACAGTGCAGGGCA
AAATATTCAGTGGCCAATTCAATACAAACAATCTCTTAAATGGGTTTCATGATGCAGTCTCCTCTTTAAACAAAAACAAACA
AAACAAACTATACTTGAACAAAAGGGTCAGAGGACCTGTATTTAAGCAAACTAGCAAAAAGTGGGGCAGAGCTCCCAAGG
25 AGAACAATATTCAGAATATTCATATTGGAATAACACAATTTTAAATGGCAGCAGAAAAGTGTGTGAAATTTCTTGATTTG
AGTTGATTGAGAAGAGGACATTTGGAGATGCCATCCTCTTCTCTCTTCTCGTTTGTCTACTACTAGTATGAGTACACATTTAA
GGATGGGGTTATGAACCTTCTGTAGCTTTATGGTCCTAAAGCAAAATAAAACTATTGCAATGAAAGACAAGAAAATCAGG
TATTAATCTTGGATAGCTAATAATGAGCTATTAATACTAGCCTGGGACAGTTTATCATGAAGCCTGTGGATGATCAATCCTTT
ATTATTATTTTTTTTTTTTGA AAAAGCTCATTTTCATGCTCTGCAAAAAGGAGAGACTCCCATGAAGCCTTTTGAAGGGATCAT
30 CATGCAGCTCAACTTTCTGTTGGATTCCATGCTAAGCAAGCTAACCTTATCCTGCATTGTTAGCACTAGGCACCCAGCTGCCAC
CTCTCCATCCTGCTGCCCTTAGGCCACATGGGAGCAGTCCATGCATGACAGCCTCTATCCTACAAGGCCTATGAGTATGGATTG
GGGGGGCCAAAAGGAAAAGCTCCATGTGCCTCTTTGTCTGCGTGGGTGAGAAGAGTTGTGCACGCAGATTAGCAGGCCAAGGT
CTGAGCCACAGCAGCATTTTTATTTTCAGATTTTGATAACTGTTTATATGTGTTGAAAACCAAAATGACATCTTTTTAAAGCTTA
TCCATAAAAAAATAGATGTCTTTTATAGTGCAAAAACATGGGGAAAAAATCATCTATTTTGATGCAGCATTTGATAATG
35 ATAAACACCTCACACCTCACTCTTTATAGTGCAAAAATGAATGAGGTCTGGGCTAGGTAGAAAAGGGTCAATGCTATTTTT
GTTTTAGAAATCATTACCTTTTACCAGCTTTTAAACCATCTGATATCTATAGTAGACACACTATCATAGTTAACATAGTTAAGTT
CAGCACTGTCTCATTTTTAATGTAAGATTGCTTCCATTTTCTACAGGCAGTCTCTCTCTTCTCCTCAGTCCCAGTGTGCAG
GTGCTATTGTTACTCTTACGAATATTTTCAGTAATGTTATTTCTTCTAAGTGAATTTCTAGCCTGCACCTTTGATGTCATGTG
TTCCCTTTGTCTTTCAAACCCAAGGTTCCTCTGTGGCCCTCTCCCTTACCCTGGGAAGGCCTCTTGGAGACCTTACCCTTGGC
40 TGTTTGGACTTTGTATACTTTAAATAATTTAACTACCCTTAATTACTTAAAAA AAAAAAAAAA

121

MNSNFITFDLKMSLLPSNLSFAFITLCPGAIFFLPDSSKLLSGVLPHSSPALQPAADHKPGPGARAEDAEGRARRREBAGPD
PEAALEDNLARIRENHERALREAKETLQKLPEEIQRDILLEKKKVQDQLRDKAPFRGLPPVDFVPPIGVESREPADAAIREKR
AKIKEMMKHAWNYYKYAWGLNELKPI SKGGHSSSLFGNIKGATIVDALDTLFIEMEKHEPPEAKSWVEENLDFNVNAEISVFE
45 VNIRFVGGLLSAYYLSGEEIFRKKAVELGVKLLPAFHTPSGIPWALLNMKSGIGRNWPWASGSSILA EFTLHLEFHMHLSHLS
GNPIFAEKVMNIRTVLNLEKPKQGLYPNYLNPSSGQWQHVS VGLGDSFYEYLLKAWLMSDKTDLEAKMYFDVAQAIETHL
IRKSSSGLTYIAEWKRGLLLEHKMGLTCTFAGGMFALGADAAPEGMAQHYLELGAELIARTCHESYNRTFMKLGPEAFRFDGGVEA
IATRQNEKYIILRPEVMETMYMWRLTHDPKYRKWAVEAVEALENHCVRNGGYSGLRDVYLLHESYDDVQSSFFLAETLKLYLYL
IFSDDDLLPLEHWIFNSEAHLPLPKDKKEVEIRE

122

CCAACTTATTTAAACAAAACAATTTTGTAGGTATTATTATACCCATTTTACAGATGATGATAAATGAGACCAATAGAAGTTAA
ATAACTTGCCAAAGGCCACACAGCTGGTGACTGATGGAGAACGAATTTAAACTCAAGTGAGCATAATTCTAAAAGCCATCTTCT
CGTTAGTGTTTCTCACTATCCAGGTCTGCCCTTTGCCCTTATTTAACTGAAGTTAAGCCATCCTTACCTGTGATCACCTAGCCTCT
5 CAGTTTGGGGGGATCATTACAGCGGGTTTTAACTCCCAATGTTCTGGTCCAGTTTGTCTTACATGTTCTTATTATACATTGT
CAAGGATGACCTCAGGACAGTACAGCAAGGACACAGTGGCACTTTCACATTTTGTTCACGAAATGACTGGGGCATAATCTCAG
ATCATCTTCTTTAGAAATGTGGAACATCAGCAGAAGAATATTAGTCTTTATACAAGTCAAATCCAAAATGACACATGTGAAAA
CTAATAGAGCTGACTTTCAGCCATGATAGCTTTGGCACACCTCACATCCCTTTGTTCAACCTCTCTTCCCTCAACGGAGAGCTG
CATTCTGGGAATTTCTGTTGTGCACTTTTCCCACTTGCCTGCTGTCTATTAAAGGTGAACATTTCTAGTTTGTCTAAGAAAA
10 CTTTTCCTTCATTTGGAATGAACAGCAATTTTATTACTTTTGACCTTAAATAGATTGTCTGCCTTCAAATCTTTTCAGCGCCT
TCATCAGCTCTGCTTCGGGGCGATCTTCTTCTGCGCAGCTCCTCCAAGCTGCTCAGCGGGTCTGTTCCACTCCAGCCCCG
CCTTGCAGCGCGCCGCGACCAAGCCCGGGCCGGGGCGCGCCGAGGACGCGGCCGAGGGCGAGCCCGCGCCGCGAGG
AGGGGGCACCCGGGACCCGAGGCGCCCTGGAGGACAATTTGGCCAGGATCCGCGAAAACCAGAGCGGGCTCTCAGGGAAG
CCAAGGAGACCTTGCAAGCTGCCGAGGAGATCCAAAGAGACATCTTACTGGAGAAGAAGAAGTGGCCAGGACAGCTGC
15 GTGACAAGCGCCGTTTCAAGGCTGCCCCGGTGGACTTCTGCCCCCAATCGGGGTGGAGAGCGGGAGCCCGCGACGCCG
CCATCCGCGAGAAAAGGGCAAAGATCAAAGAGATGATGAAACATGCTTGGAAATAATTATAAAGGTATGCTGCGGATTAATG
AACTCAAACCTATATCAAAGGAGGCCATTCAGCAGTTTGTTTGGTAACATCAAAGGAGCAACTATAGTAGATGCCCTGGATA
CACTTTTATTTATGGAATGAAACATGAATTTGAAGAAGCAAAATCATGGGTGGAAGAAATTTAGATTTTAAATGTGAATGCTG
AAATTTCTGTCTTTGAAGTAAATATACGCTTTGTTGGTGGACTACTCTCAGCCTACTATCTGTCTGGAGAAGAGATTTTTCGAA
20 AGAAAGCAGTGGAATTTGGGTAAAATTGCTACCTGCATTTTCACTCCTCTGGAATACCTTGGGCATGCTGAAATATGAAAA
GTGGTATTGGAAGGAATGGCCCTGGGCCCTGGAGGCAGCAGTATCTTGGCAGAATTTGGAACCTGCAATTTGGAGTTTATGC
ACTTGAGCCACTTATCAGGAAACCCATCTTTGCTGAAAAGGTAATGAATATTCGAACAGTACTGAACAACTGAAAAACCCAC
AAGGCCCTTATCTTAATATCTGAATCCAGTAGTGACAGTGGGGTCAACATCATGTATCAGTTGGAGGACTTGGAGACAGCT
TCTATGAGTATTGCTGAAGCCCTGGTTAATGTCTGACAAGACAGATCTGGAAGCTAAGAAGATGTATTTTGATGCTGTTTCAGG
25 CTATCGAGACTCATTTGATCCGCAAGTCTAGCAGCGGACTAACTTATATCGCAGAGTGGAAGGGGCTCCTGGAGCACAGA
TGGGCCACCTGACCTGCTTCGCGGGGGGCTGTTGCGACTCGGGGCTGATGCAGCTCCCGAAGGCATGGCCCAACACTACCTTG
AACTCGGGGCTGAAATGCCCCGCTACTTGTCTGATGAATCATATAATCGAATTTATGAACTGGGACCAGAAGCTTTCAGATTTG
ATGGTGGTGTGTAAGCCATCGCTACAGACAAAATGAAAAATACATCTTACGCGCCAGAAGTTATGAGAGCTTACATGTATA
TGTGGAGACTGACTCATGATCCAAAGTACAGGAAATGGGCTGGGAAGCCGTAGAGGCCTTGGAAAACCATTCGACAGTGAATG
30 GAGGCTATTCAGGCCATAAGGATGTTTACCTTCTTCATGAGAGTTATGATGATGTGCAGCAGAGTTTCTTCTGCGAGAGACAT
TGAAATATTTGTACCTAATATTTTCTGACGACGATCTTCTTCCACTGGAGCATTTGATCTTCAATAGCGAGGCACATCTTCTCC
CTATCTCCCTAAAGATAAAAAGGAAGTTGAAATCAGAGAGGAATAAAAAGACATTTATATTTTATTCTGCTCCATTCCTCTC
ACTGTATACCTTAATAATTCCTTTTCTGGTAATCAGGCACATGATGAACTTTGATTAGTAGGTCTGTGATTAAGTTCTTAAAT
GTTTTGCAGTCTTTTATGTTTATTTATCATAGGTATAGGTGGACCTAAATTCCTTATCATATCTTTTATTAATTCAGCCAGGTAT
35 CCACAGTTTTTTGTATTATGTTTAAAGTAACCTATTTCTCTGGATTTCATGAAGGTGTAATATCGTTTTTGTAAACGAAT
AGAATGTATAGCGATGACCTCTTAATTTAATTTGATTGACTGCAAACTTTTCTCTCTTAAGAGGAGATGATGTCTGCT
TTAAGCTGTAATGTTTGGCATGTTGCAAAAAGCCATAATTAAGTATAAAAAAGCTTTTCTCTTACAATTTTCATGTTAATC
TGGTTGTCTGTCCACCAGAGACAGATCTTCTGTGACAGCTCCTTATGCAGGTCTATCATTTTGTATAGAATGTCTTCTAAA
ATACTTCACTCACATTTGTAATTCAAATTAGAAAGTCATTTCAAAAAGGTGATGTCATGTTGACCTCATTTTCATCGGAAGTCAGT
40 ATATTTTGTGTTAATTTATATTAGTGTCTTCTATTTTGAATAAAAAAAAAAAAAA

124

TTTTGGTCTGAGAATGATGGACATTTAGACACTGGCGCCAGGTTTGCCTGGACCGGCGCAGCAGGGGTGGGGCGACGAAAG
ACACACAGGTGGGCTACAGGTGTCACACGGCACCAGGCCAGGCCCCGGGTGGCTGGGGTGAGGATGGGTGTTTGGCCAGTGACC
AGGAGTCAGGTCAAGTCCAGGTGGTCAAGTCCAGGGGTCCAGGAGGGAGGAGTGCATTAACCTCTGCTGTCACGCGTC
45 ACCAGGCGGTGCTCACAGAAAGCAACCTCGGCCCGGGGCCG

126

AATTATATATTACATGTTTATTAAGAGCACAACTTTTATGTAAATTTACATTTAATGAAAAAATCAAAAATATTTACAAAAT
CTTGAAGACAGATGTGATGTTCTAATTAACAATCCAAAGTAGTAAATAACAATCCTTTAAAACCTCACATTTATTAGAGTTGT
GTTTACAAATCTTGGTTAAAGAGGCAGCTACAAAGTTTACCACATATATAAGCAAGAACCAGCTTGCTAGGTACATTTCCCA

TTGAAAATCTACTGGTCTCTTTTACACCATTAGTGGATTTTAAATGGAAAAAATCAATATAAACTCATATGGCTTCAAA
ATTGTA

128

- 5 AATATAGTTATCTTCTTAAAAACCATTTATAACAATTCAGAGAGAGTTCTTTACAAAGCCATGAATATGAACTATGGGAATCAT
GGTTCCTTTTAAAGCAATTTTCAAAATAAGTACCAATTAAGCTTTAGGTTCCAAGAAGATTCTGGGACTCAGGAAGAAAAAGTG
CCATCAGGTGACCAGCTGTTGCATTTCTTGCTTATTTCTGTTTGTGTTTGCACATCATAATGGATTTTCTTAGTGCCCTAATT
GTGAAGGGTTTCTCTAGCTTTGGTTATGTGTAATGTTTCACGTGACCTTTTGTCAATCGTTTGGAAATTTTCTTTCTTTTC
TGTGCTTTATTACTAATAAGTCCAATGAGTGAGTAGCTAGATGACTAGTATGTAGTTTATATTTTGGTAAAAATATTATTGC
CCTTTCAGAAATGCCTCATCTAAAGATACATGATAATTTTGGAGTTGGAGGGGGCCTTAGAGGCTCTCCAGCTCTGCTTCTTGC
10 CCATTGCCAAATACTGAAATGGAAGCCCGTCTTACCTGGGGTCACTAACTGGTTGGTTAACTGAGCTAAGAATAGGCTGTGGGT
CTCCTCACTTGTGGCCAGTGCCTTTCTGCTATACAAAATGTCTAATCTCAGATTTTCTTCTGCTGCTTACTGCTTCACT
GGATGAACTACAAAAACCCATGATTAAAGGTTTATGAATTCAAGTAATAATTAGATTTTTTGCACAGACTTACTTAACTTCC
TTATTGGATATGTTTGTAAACACATAAACACAAAGCACTTTTCAAACATGATGCACTTTTATCTTTGTGAATAATTTACTGTCTT
TTCTCTCTGGGATATGAGAAACATTTTAAAAACGTTTAAACAGAAGAGAGCAATAAAGATATATCAGGAAGGATGTATTAG
15 TTATTTACTTAAATGTTTATAATATCTGGATTTTTTGTGTTTGTACTCATAGAACTGGTGTGTTGCTGTTTTTATTTCTC
TAATTGTTGCAGAGTTCTGCCCTGTACAAAGCTACAGAACTGATTTGTTTTATTTCTTCTGAGCACATGTTAAACAACTA
AGCTTCACATTAGAGTGATGTCATAATGTAAATGTTTGCATTTGTGGTTAGGTATTGAAGTTTATGTCTGTCTGTGTAAAGAT
TCATCTTTTATTGTAAATATTAGACTTTACCACAGAAATATTGGAACAGTTTGTCTTTATAAGATTAAAAAGCATCTTCAGAA
TGGAGCTTGCCTTGTGCTTAGAAATAATATGTTGAACATTTTGAATATACTATTTTAAATCTAAATCTGTCACTTCGCTGC
20 CTCTTAAAAATAGTGTGGTATTTCAAATATTGCTAGAGCTATTTCTGAAATACATTTGCAAAATAAGGCTGCTTTGTAATCA
AGGAATATTTTATTGATTGAAGGAAATGACTGTACTGCGATTCAAAAGTAAACTTATTTTATTATACAGATTATTTCTTAAAA
ACTCTATTTATACCTTAACATGAAATCCATGACCACACCAAACTTGGTTATTCATAATTTTCTGTTAAATATAAAACACTGT
AAGTTAAAAACAGTAATGCCAATTTGAATTTATTTTGGAGTCAAGAACCAGTTGTTCTTTTATATTAGATGAGGATGAT
TGAGTCCATATACTATGTATGTTTACATATACTATACATGCACATTAGGTGTTTTCATTTGTGTTTTGCTTATGAAATGTCATT
25 TAAAGTTCACCTCTTGAGCATCAATAAAAAGGAAGCTGTGTGGTTTTGG

129

MSLMVSAGRGLGAVWSPHVVQLVQLARGLRAKPGGTS DAYAVIQVQKEKYATSVSERSLGAPVWREETFELPSLLSSGPAA
AATLQLTVLHRALLGLDKFLGRAEVDLRDLHSSLGKSFFKTLKKRAWAIFLRLCLKKN

130

- 30 GTGGAGGCCCGCAGTCGCGCGCATCTTCTCTCGCTTCTGGAGTGTATCGTCACCATGTCCCTAATGCTCTCGGCTGGCCGGG
GCCTGGGGGCCGTGTGGTCCCCAACCCACGTGCAGGTGACGGTGCTGCAGGCGCGGGCCCTGCGGGCAAGGGCCCCGGGGCA
CGAGCGACGCGTACGCGGTGATCCAGGTGGGCAAGGAGAAGTACGCCACCTCCGTGTCGGAGCGCAGCTGGGCGCGCCCGTGT
GGCGCGAGGAGGCCACCTTCGAGCTGCCATCGCTGCTGTCTCCGACCCGCGCGCGCCGCCACCTTCAGCTCACCGTGTCTGC
ACCGCGCGCTGCTCGGCCCTCGACAAGTTCTTGGGCGCGCGAGGTGGACCTGCGGGATCTGCACTCCAGCTGGGCAAGAGTT
35 TTTTAAAGACTCTTAAAAAAGAGCCTGGGCAATTTTTTAAAGACTCTGTCTTAAAAAAACTAAAAAGAAAAAAGAGCCCC
TTCACCTACAGGGGACAGGAGACCATGGATTGGACCCAAAGGATTGAATGCATCTGCATGCTGTCTTTGAACACTTTTC
TCTCCCTGCCAAAAGGAAACCCAAATATTGTTGGGATACTGGGGAATTTAGTAGAAGGGCTTAATGTAGTTAATAAAGTT
AAAAGTCAGTAGAAAACAGGTGCCTCAGCCTTCAAATGGTTGCTTTTTTCCATTTTCCCTCATGAATAGACTCACCAGCATTT
TACCCCTTGTATATAAACTGTGCAGAGCAAGAAGATGATACTTATTTTGAATTTGTATTTTAAACTAGATTATAGACTTT
40 TTTTTTTTTTAACTAGGGCACTTGCCTTCTTCTAGCTAAAAGCACCAGCTGAGATTTTTCAGGTAAATTTGTGTTACTCAC
TTAAGACTGGAATTAGAATGTTTCTGGTATTTCTCATTTTTTCCCTGGCTATGATAGAATCTCATCTAATCTTGACATCTCTC
CTAGGGGAAGAATATCACAGGCTAATAGCGTGGTTGGGGGTGAAGATGATAGCAGTTATTAATCAGGAATCTCTTTTATGTAT
GTCCTTGTATACATTGAGGTTAAGAGACAAAATCATTTGGCAGTGCAATCTCTTTCCAGGATTTCTGTTGCTGTGGCATTTGGTTAT
ATCAGAGCACTTTAATCTGAAGGATGATACTGTAACCTGATTATCTAATTAGCTTTTAAATATCTACAGCTATTTTATTTTAT
45 TTAATTCCTCTACAGTACTGGGACCACTGTAACTTCTCAGATGACTGTATTTTGTAGTGTCTATGAAATTTATTTACATA
CTTATAAAGAAACTGTATGTTCACTTGAAAAAATTTTTTTTTT

131

MCAPFWLLLLLLLEQGSQRRWRWCGSEEVAVLQESISLPLEIPPDEEVENI IWSSHKSLATVVPKGEGHPATIMVTNPHYQG
QVSFLDPSYSLHISNLNLSWEDSGLYQAQVNLRTSQISTMQYNLCVYRWLSEPOITVNFESSGEGACSMVLCSVEKAGMDMTYS
WLSRGDSTYTFHEGVLSTSWRPGDSALSYTCRANNPISNVSSCPIPDGPFYADPNYASEKPSTAFCLLAKGLLIFLLLVILAM
GLWVIRVQKRHKMPRMKKLMNRNMLRKEAKPGSSPA

5 132

ACTAGGAGCCTCCTAATGCAGTCTTCTGCACAGTCTTGGGACTGACTGACTGAATCACACCTCTGGGGCTGGGGGCTGCTGAC
ATGTGTGCCTTTCTTGGCTGCTTCTTCTCTGCTGCTCCAGGAGGGCAGCCAAAGGAGACTCTGGAGATGGTGTGGATCCGAG
GAAGTGGTTGCGGTCTTTCAGGAGTCCATCAGCCTCCCCCTGGAAATACCACCAGATGAAGAGGTTGAGAACATCATCTGGTC
TCTCACA AAAAGTCTTGCCACTGTGGTGCCAGGGAAAGAGGGACATCCAGCTACCATCATGGTGACCAATCCACACTACCAGGGC
10 CAAGTGAGCTTCTCGACCCAGCTATTCCCTGCATATCAGCAATCTGAGCTGGGAGGATTGAGGGCTTTACCAAGCTCAAGTC
AACCTGAGAACATCCCAGATCTCTACCATGCAGCAGTACAATCTATGTGTCTACCGATGGCTGTGAGAGCCCCAGATCACTGTG
AACTTTGAGAGTTCTGGGGAAGGTGCCTGCAGTATGTCCCTGGTGTGCTCTGTGGAGAAGGCAGGCATGGATATGACCTACAGC
TGGCTCTCCCCGGGGGATAGCACTTATACATTCATGAAGGCCCTGTCTCTAGCACATCTCTGGAGGCCGGGGGACAGTGCCTCT
TCCTACACCTGCAGAGCCAACAACCCCATCAGCAACGTCAGTTCTTTGCCCCATCCCTGATGGGCCCTTCTATGCAGATCCTAAC
15 TATGCTTCTGAGAAGCCTTCAACAGCCTTCTGCCTCTTGGCCAAGGGATTGCTCATCTTCTTGTCTCTGGTAATCTTGGCCATG
GGACTCTGGGCTCATCCGAGTCCAGAAAAGACAAAAATGCCAAGGATGAAGAACTCATGAGAAACAGAAATGAAATTGAGGAAG
GAGGCAAGAGCTTGCTCAGCCCTGCCTGACTGCTCCTTGGGAACCCGAGTCTGAGCTTGGTTTCTTCCAGCACCCAGAGAA
TCCTTCTCTCAGCTCTCTTCTTCCAGGGGAAGGAGGTGCTCAGGGGTGGGTATCCAGAGAGCCATACTTCTGAGGGAAGACTGG
CTGGCAATAAAGTCAAATTAAGTGACCACCAAAAAAAAAAAAAAAAAA

20 133

MIEEKSDIETLDIPEPPNPSYECQLRLRLSTGKDLKLVVRSTDTVFHMKRRLHAAEGVEPGSQRWFFSGRPLTDKMKFEELKI
PKDYVVQVIVSQPVQNPTPVEN

134

25 ATATCAGCTTCCAGTGTATGTGCTTGGCACCGCCAATCAACATGATAGAGGAAAAGAGCGACATAGAGACTCTGGATATTCCTGA
GCCACCACCCAATTCTGGATATGAATGTGAGCTTCGTTTTCGCGCTTTCCACAGGCCAAAGACCTCAAGCTTGTGGTTCGCAGCAC
AGACACAGTATTTCCACATGAAGAGACGGTTGCATGCAGCAGAGGGAGTGGAAACAGGTAGTCAGCGGTGGTTTTTTTCTGGCAG
ACCTCTCACTGACAAAATGAAGTTCGAAGAGCTGAAGATCCCAGGACTATGTTGTACAGGTTATAGTGAGCCAACCTGTGCA
30 GAACCCAAACACAGTGGAGAACTGAAGTGAAGCCCTGTTGGCCAGCTCCCACATCCCTCTGCTCCTTTTTATGGTCTCTGTTGT
ATTTCTACTCTGCGCGGTGAATCTATTTCACTGCTCTAAATTCCTATGAATGGATTTAGTTCGAGGAATTACCACTGAAA
AATTCCATCTGTGATGGAGACCAACAAAATAATAAACACAAAGAGCCAGGCTTTGAGACTCATGTAATTACAATTTCTAATT
TGAAAGGCAGTTAAGAAATAGATAACCATTTATTTTAGAACACTCAACAACATATGTAATGGCTATATTTCACTGACTTGGACTG
TAAATGAAACATTTGCATCCATGAAGGACAGCACCAAGCACCCTTTTGAAATACAGAATTTTTTTAGAAAAATATATCAAAATATA
TAATTTCCAGAAACCATAAAATAATGGATATAAACTTAACTTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTT
TAATGAAACTGAAAAGAGCAGTATTTGAGGTTGCTTCTATTCTGGTTTTTTATTTCTTAGCTCAATTAATATTTAGCCATAAAT
35 GAGTAGAGACTGGCAAATGTGCTTTAGTGTGCTTTCTCATCCCCACATCTTGAGCTCCTTATTTACATTCACACTAAATTT
TGGTGCCCTTCAGCACATTAGTGGCAGGCACCCCTCTGGAACACTAGGCAATAATTTCAATCAATACAGTCAGGTCTCTTGAGTT
TCAACAGATACTCAGTTGAAAAGTCGCTGTCTGCTGCATAAGTATTTTGAAAGGTCTGTATAACGAAGCCATTTTTATAT
CCAGGCTTAGAAGGTCACACTATATAGTACCTTCATTGATCTATCTATTTCTGCTTGGAACTTTTCATAGCTAAGTATAACCCC
CAAATGCATGGTTCTCGGGTCAGGAGACACCAAAATCAATTATAGCTGTTCCCACTCAAGTTAATAAAGTAAATGATTTCTCTC
40 ACTTTGCATGGAGGGGTGAAGGAAAGCCTATTTTATCTGTGCTGGGAGAACTGTGCCATTTTTCAGTCTTTTAGAGGAAATTT
TCAACTCAAAATTTTTAGTGTGAAAGGTTTACTGGTGTACATAAACATTTAGCAGTGAGACCAATAATGAAACATTGCTTTTA
TACCTTAGTGCTTTCCAGTTCACTGTTACCTCTAGTCATGGTGTAGATGACAATTTTCTCCCTCATCTTTTGTAGCAAAGAAGCA
AATTAAGAGCCTTTTGCAAGTTAGATTAGGGGACTGAAAAATCCAGGGAGATAACCTGATTTCTTATACTATGACTGGAGTAA
AGCTGCGCAAGAGTTAGGTTGAATTTGGAGCCTTTTAAAGTTGGTTCCATAAGTGAGAAGGGGGCTAACATGTGACCACTGTG
45 CCCCTGGAATCCCATGAGATGGAAAAAGTACAGACAAGAAGGCATTAACGCTGTTCTCAGTTGGAAAAATGTCAGCCCTCCACAC
CTGTTTGGAGTGTGAAGTGCATTATTAATTTTAAATGACAAATGAAACTGCTAATTTAGTTTAAATGTAAATTTTAAATGGATTTG
AATTTCTCTATTGCTGATTAATAATATCTGTAATTTGTAACATCACTGTTTCTGAGAGCTTTAACTACTTCTTATTTTATGCAA
TCATTTTAAAGATTTGTGGCTATATCTAATTTAAATTTGCTGTAGATTGAATCTAAATTCAGTGAGACCCACAGCATTAATCC

- AAAGGAGATTTTTTATGTTTAGAGGTGGAGATATTAGGGTAAATCAGTATATTGCTCCTGTTCATCAAAGTTTTTGTTCATTT
GTGTGCCACATCGAATAGCAGCACTTCTCAAAGTTTGTCAACTGGAAAAAGTGTAAAACTGTTTTTCTTAAGCAGATTT
GGTGATGTCATCAATATCTATGATATCGATTGGTAAAGGCTTTCTTTGTTCAGGATTTGAGCTCTCTTTTTAAGAGTTTATT
TTATTTTAAAGGAACATATGCTTTGAATAACAATTCTCTGGGCATGTTTAACTAACCTAGGTGGGAAAACCATGGTGCCTTTTA
5 ATTGTAAATAGATTGATATAAGATTGGACCAGTACTTGATGTGTATAATTTTAGTATGTAGGGTTTATGTGCTTTTTTAAAAA
AAAAAAAAAGGTTCAATTTT
- 136
- GGAAGGTATTAGACAAACATGCCCATGAAACCCAGCTTCCCTTTTCTATGTGCTAGGCATGGAACTTATGAAATTTTAGCAC
TCCAAAGTCATTTGGACTTCAAGGCATTTAAATCATTTTTTCTAAGGATTTAAACAGCTCCACTATAAGTCTTACCTGACATG
10 AATTGGTGAGAGACCAGGCTGATCCTGGCAAAGGCTTGTGTCTTTCTGCCAGGCAAAATCCTGGGTCTTCTAGCAGGACCTA
AGCCAGTCTGGGGACGCTGATATTGAGGATGAGCTGGGGGACTCTGCTCTGTCTCTGTGAACACACAGGAGGCCCATCCAGAG
TGAGTGAGGTTGATTCTCTCTCCCTCTTTGC
- 137
- MGTSLSPNDPWPLNPLSIQQTLLLLSVLATVHVGRLLRQRRRQLRSAPPGPFAWPLIGNAAAVGQAAHLSFARLARRYGDV
15 FQIRLGSCPIVVLNGERAIHQALVQQGSADFAPAFASFRVVSGRSMAGHYSEHWKVRRAAHSMMRNFFTRQPRSRQVLEG
HVLSEARELVALLVRGSADGAFLDPRPLTVVAVANVMSAVCFGRYSHDDPEFRELLSHNEEFGRTVGAGSLVDVMPWLQYFPN
PVRTVPREFEQLNRNFSNFILDKFLRHCESLRPGAAPRDMMDAFILSAEKKAAGDSHGGGARLDLENVPATITDIFGASQDTLS
TALQWLLLLFTRYPDVQTRVQAEQDQVVRDRPLPCMGDQPNLPVLAFLYEAMRFSFVPVTI PHATTANTSVLGYHIPKDTVV
FVNQWSVNHDVVKWPNPENFDPARFLDKDGLINKDLTSRVMI FSVGKRRCIGBELSKMQLFLFISILAHQCDFRANPNPAKMN
20 PSYGLTIKPKSFKVNVTLRESMELLD SAVQNLQAKETCQ
- 138
- ACTCTGGAGTGGGAGTGGGAGCGAGCGCTTCTGCGACTCCAGTTGTGAGAGCCGCAAGGGCATGGGAATTGACGCCACTCACCG
ACCCCCAGTCTCAATCTCAACGCTGTGAGGAAACCTCGACTTTTGCCAGGTCCCCAAGGGCAGCGGGGCTCGGCAGCGCAGGCAC
CCTTCTCCGTCCTCCATCCCAATCCAAGCGCTCCTGGCACTGACGACGCCAAGAGACTCGAGTGGGAGTTAAAGCTTCCAGTGAG
25 GGCAGCAGGTGTCCAGGCCGGGCTCGGGTTTCTGTGTGACGCTTGGCCCTAGGCAAAAGTCCAGTTCTTCTCGGAGCCGGC
TGTCCTCCGCGCCACTGGAAACCGCACCTCCCCGAGCATGGGCACCGCTCAGCCCGAACGACCTTGGCCGCTAAACCGCTG
TCCATCCAGCAGACCAAGCTCCTGCTACTCCTGTGCGTGCTGCGCCACTGATCGGAAACGCGGCGCGGTGGGCCAGGCGGCTCACCTC
CGGCAGCTCCGCTCCGCGCCCCCGGGCCGTTTTCGCTGGCCACTGATCGGAAACGCGGCGCGGTGGGCCAGGCGGCTCACCTC
TCGTTGCTGCGCTGGCGCGGCGCTACGGCGAGCTTTTCCAGATCCGCTGGGCAGCTGCCCATAGTGGTGTGAATGGCGAG
30 CGCGCCATCCACAGGCCCTGGTGACGAGGGCTCGGCCTTCGCCGACCGGCCGGCTTCGCTCTCCTTCCGTGTGGTGTCCGGC
GGCCGAGCATGGCTTTCGGCCACTACTCGGAGCACTGGAAGGTGACGCGCGCGAGCCACAGCATGATGCGCAACTTCTTC
ACGCGCGAGCCGCGCAGCCGCCAAGTCTCGAGGGCCAGTGTGAGCGAGGCGCGAGCTGGTGGCGCTGTGGTGGCGGGC
AGCGCGGACGGCGCTTCTTCGACCCGAGGCCGCTGACCGTGTGGCCGTGGCCAACGTCATGAGTGGCGTGTGTTTCGGCTGC
CGCTACAGCCACGACACCCCGAGTTCCGTGAGCTGCTCAGCCACAACGAAGAGTTCGGGCGCAGGTGGGCGCGGGCAGCCTG
35 GTGGACGTGATGCCCTGGCTGCAGTACTTCCCCAACCCGCTGCGCACCGTTTTCGCGCAATTCGAGCAGCTCAACCGCAACTTC
AGCAACTTCATCTGGACAAGTTCTTGAGGCACTGCGAAAGCCTTCGGCCCCGGGGCGCCCCCGGACATGATGGACGCGCTTT
ATCCTCTCTGCGGAAAGAAGGCGGGCGGGGACTCGCACGGTGGTGGCGCGCGGTGGATTGGAGAAGCTACCGGCCACTATC
ACTGACATCTTCGGCGCCAGCCAGGACACCTGTCCACCGCGCTGAGTGGCTGTCTCTCTCTTCCACAGGTATCTGTATGTG
CAGACTCGAGTGCAGGCAGAAATTGGATCAGGTGCTGGGGAGGGACCGTCTGCCCTGTATGGGTGACAGCCCAACCTGCCCTAT
40 GTCTTGGCTTCTCTTATGAAGCCATGCGCTTCTCCAGCTTTGTGCTGTCTACTATCTCTCATGCCACCACTGCCAACACCTCT
GTCTTGGGCTACCACATTTCCAAGGACACTGTGGTTTTTGTCAACAGTGGTCTGTGAATCATGACCCAGTGAAGTGGCCTAAC
CCGGAGAATTTGATCCAGCTCGATTCTTGGACAAGGATGGCTCATCAACAAGGACCTGACAGCAGAGTGTATGATTTTTTCA
GTGGGCAAAAGCGGTGATTTGGCGAAGAACTTTCTAAGATGACAGTTTTTCTCTTCTCATCTCCATCTCTGCTCAGGAGTGCAT
TTCAGGGCCCAACCAATGAGCCTGCGAAATGAATTTTTCAGTTATGGTCTAACCATTAAACCAAGTCAATTTAAAGTCAATGTG
45 ACTCTCAGAGAGTCCATGGAGCTCCTTGATAGTGCTGTCCAAATTTTACAAGCCAAGGAACTTGCCCAATAAGAAGCAAGAGGC
AAGCTGAAATTTAGAAATATTCACATCTTCGGAGATGAGGAGTAAAAATTCAGTTTTTTTTTCCAGTTCTCTTTTGTGCTGCTTC
TCAATTAGCGTTTAAAGTGAGCATAATCAACTGTCCATCAGGTGAGGTGTGCTCCATACCCAGCGGTTCTTCATGAGTGTGG
GCTATGCAGGAGCTTCTGGGAGATTTTTTGTAGTCAAAGACTTAAAGGCCCAATGAATTATATATACATACTGCATCTTGGT
TATTTCTGAAGGTAGCATTTCTTGGAGTTAAATGCACATATAGACACATACACCCAAACACTTACACCAAACTACTGAATGAA

AACTGTGGA AAAATGAGACGGTTGCTTGAACATTTGTTGGCTTGGCTGAAGAGGAAATAAAAACAGAACAGGAGGTGGTAGAGG
GCATGGATATCTCTACTCGCTCCAAAGATCCTGGCTCTGCAGAGAGAACAGCCCAGAAAAGAAAGTTCCCCAGCCCTCCACATT
CTTCCAATGGCCACTCGCCGAGGACACATCAACAAGCCCCATTAAAAAGAAAAGAAACCTGGCTTACTGAACAGTAACAATA
AGGAGCAGTCAGAACTAAGACATGGTCCGTTTACTATATGAAGCAGCCACTCACCACAGACCCCTGTGTGATGTTGTACCGCAGG
5 ATGGACGGAATGATTTCTACTGCTGGGTTTGTACCGGGAAGGCCAAGTCCTTTGCTGTGAGCTCTGTCCCGGGTTTATCAG
CTAAGTGTCTGAGACTGACATCGGAACCAGAGGGGGACTGGTTTGTCTGAATGTGAGAAAATTACAGTAGCAGAAATGCATCG
AGACCAGAGTAAAGCCATGACAATGCTCACCATTGAACAGTTATCCTACCTGCTCAAGTTTGGCCATTGAGAAAATGAAACAGC
CAGGGACAGATGCATTTCCAGAAGCCCTTCCATTGGAACAGCACCCCTGACTATGCGGAATACATCTTCCATCCAATGGACCTTT
GTACATTTGAAAAGAAATGCGAAAAAGAAAATGTATGGCTGCACAGAAGCCTTCTGGCTGATGCAAAAGTGGATTTTGCACAACT
10 GCATCATTTATAATGGGGGAAATCACAATTTAGCGCAAATAGCGAAAGTAGTCATCAAAATCTGTGAACATGAGATGAATGAAA
TCGAAGTATGTCCAGAATGTTATCTAGCTGCTTGCCAAAAACGAGATAACTGGTTTGTGAGCCTTGTAGCAATCCACATCCTT
TGGTCTGGGCCAACTGAAGGGGTTTCCATTCTGGCTGCAAAAGCTCTAAGGGATAAAGACGGCAGGTCGATGCCCGATTCT
TTGGACAACATGCAGGGCCTGGGTTCCAATAAATAATGTCTACCTCATGTCTAAAGAAATTCCTTTTCTGTGAAAAAGACTA
AGAGCATCTTCAACAGTGCCATGCAAGAGATGGAGTTTACGTGGAGAACAATCCGCAGGAAGTTTGGGGTTTAAATTACTCTC
15 CATTTAGGACACCCCTACACACCCCAACAGCCAGTATCAAAATGCTGCTCGATCCCACCAACCCAGCCGCGCAGTCCGAATAG
ACAAGCAGGAGAAGGTCAAGTCAACTTTGACATGACGGCATCCCCCAAGATCCTGTAGTGAAGCAAGCTGTGTGAGGGGCA
CAGGCCGCGGATTTCTTGTGCGATATGCCGCGCTCCCCCATGAGCACAACTCTTCTGTGCACAGGGCTCCGAGCTGGAGC
AGGATGCTGAGAAGAAGGCCACGTCGAGCCACTTCACTGCGAGCGAGGAGTCCATGGACTTCTGGATAAGAGCACAGCTTCAC
CAGCCTCCACCAAGACGGGACAAGCAGGGAGTTTATCCGGCAGCCCAAGCCCTTCTCTCTCAACTGTGAGCTCCTATACGA
20 CGAAAACGGACAAAACCTCCACCACCGGCAGCATCTGAATCTTAACTGGATCGAAGCAAAGCTGAGATGGATTTGAAGGAGC
TGAGCGAGTCCGTCCAGCAACAGTCCACCCCTGTTCTCTCATCTCTCCCAAGCGCCAGATTCTGAGCAGGTTCCAGCTGAATC
TTGACAAGACCATAGAGAGTTGCAAAGCACAAATAGGCATAAATGAAATCTCGGAAGATGTCTATACGGCCGTAGAGCACAGCG
ATTCGAGGATTTCTGAGAAGTCAGATAGTAGGATAGTGTATATCAGTGATGATGAGCAGAAGTCTAAGAACGAGCCAGAAG
ACACAGAGGACAAAGAAGGTTGTGATGAGCAAAAGAGCCATCTGTGTTAAAAAAGGCCAAGCTTACAAACCCAGTGGAGA
25 TTAAAGAGGAGCTGAAAAGCACGTCACCAGCCAGCGAGAAGGCAGACCCCTGGAGCAGTCAAGGACAAGGCCAGCCCTGAGCCTG
AGAAGGACTTTTCCGAAAAGGCAAAACCTTCACCTCACCCTATAAAGGATAAACTGAAGGGAAAAGATGAGACGGATTCCCCAA
CAGTCCATTTGGGCTTGGACTCTGATTTCAGAGAGCGAATTTGTCTATAGATTTAGGAGAAGACCATTTCTGGGCGGGAGGGTCGAA
AAAATAAGAAGGAACCCAAAGAACCATCTCCCAACAGGATGTTGTAGGTAAGTCCACCATCCACGACGGTGGGCGAGCCATT
CTCCCCCGGAAACACCGGTGCTCACCCTCTTCCGCCAACTTCCGCGGCTGGCGCCACAGCCACCCAGCAGCTCTCCCA
30 CGGTACCGTACAGGCCCCGCCCCGCGCCACAGGAAGCCAGTGAAAAAGCAGAGGCGCTTTTACCGAAGGAGACTGCC
CGGCCGTGCAGCGGTCGTGTGAACTCATCAAGTAAGTTTCAACGTCCTCCCAAAAGTGGCAGATGCAGAAGATGCAGCGTC
AGCAGCAGCAGCAGCAGCAAAACCAGCAGCAGCCCTCAGTCTTCCAGGGGACGAGATATCAGACCAGCAGGCTGTGA
AAGCTGTCCAGCAGAAGGAGATCACACAGAGCCATCCACGTCACCATCACCCTGGTGACCAGCACAGTCATCGCCCCCTGG
TCACCAGCTCGGGTCCATGAGCACCTTGTGTCTCAGTCAACGCTGACCTGCCCATCGCCACTGCTCAGCTGATGTGCGCG
35 CTGATATTGCCAAGTACACTAGCAAAATGATGGATGCAATAAAGGAACAATGACAGAAATATACAACGATCTTTCTAAAAACA
CTACTGGAAGCACAATAGCTGAGATTCGCAGGCTGAGGATCGAGATAGAGAAGCTCCAGTGGCTGCACCAGCAAGAGCTCTCCG
AAATGAAACACAACCTAGAGCTGACCATGGCGGAGATGCGGCAGAGCTGGAGCAGGAGCGGGACCGGCTCATCGCCGAGGTGA
AGAAGCAGCTGGAGTTGGAGAAGCAGCAGGCGGTGGATGAGACCAAGAAGAGCAGTGGTGGCCAACTGCAAGAAGGAGGCCA
TCTTTTACTGTGTTGGAACACCAGCTACTGTGACTACCCCTGCCAGCAAGCCACTGGCCTGAGCACATGAAGTCTGCACCC
40 AGTCAGCTACTGCTCTCAGCAGGAAGCGGATGCTGAGGTGAACACAGAAACACTAAATAAGTCTCCAGGGGAGCTCTCCGA
GCACACAATCAGCACCTTCAGAAACGCCAGCGCTCCAAAGAGAAGGAGACGTCAGCTGAGAAAAGCAAGGAGAGTGGCTCGA
CCCTTGACCTTTCTGGCTCCAGAGAGACGCCCTCTCCATTTCTCTTAGGCTCCAACCAAGGCTCTGACCATTCCCGGAGTAATA
AATCCAGTTGGAGCAGCAGTATGAGAAGAGGGGATCGACAGTTCCGATCACAACACCAGTACCAGCACGAAGAGCCTCTCTCC
CGAAAGAGTCTCGGCTGGACACCTTCTGGGACTAGCAGTGAATCGGGACACAAACCACCCACCCCATTTGGGAGAAAAACCCAGA
45 CGCCAGGAAAAAGAAACAACAAGGCAGGAGAACAGCCACTTTCAGACTTGAATAAGTGAACAAACCCCTCAGTTGAGCCTGAGC
CCCCGGCGCGGGGCTGTACTACTACAGGACACCAGCATCGGCTTTGACTGCAGACTGTTTACCCACACGAGCCCTGTGCTTT
TGGTGTAAATAATGTACAAATTTGTGATGTCAATTGAATCTAGAGGACTTTCCCTTTTATATTGTATTAACTTTAACTTATT
AAAAAAGAAAAGAAAAGAAAACGATTT

141

50 MALSSQIWAACLLLLLLLLSLTSGSVFPQQTQLAELQPQDRAGARASWMPMFQRRRRRDTHFPICIFCCGCCHRSKGCMCKT

142

TCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCAGATGGCACTGAGCTCCCAGATCTGGGCCGCTTGCCCTCCTGCTCCTCC
TCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCACAACAGACGGGACAACCTGCAGAGCTGCAACCCAGGACAGAGCTG
GAGCCAGGGCCAGCTGGATGCCCCATGTTCCAGAGGCGAAGGAGCGAGACACCCACTTCCCCTATCTGCATTTTCTGCTGCGGCT
5 GCTGTCATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCGTCCTCCCTTCCTTATTTATTT
CCTGCTGCCCCAGAACATAGGCTCTGGAATAAAATGGCTGGTTCTTTTGTTTTCC

144

AAAGTCCTTGTAACATTGAGTTACAGGGCTTTAACTCCTGTGTCTGAAAAATCACAACACTGATGACAATCAAAGCCTCATCT
TAAGCCCCGTAGAAGATGCCAATCAAAATAAATGCATTCCTGAGGCACTAGGCAAGAAATTAAGCTATTCAACTCCTCAAG
10 GCCCAGGGACTATTGCGGAAGAGGTGGGCGCGTAAGATTGTAAGGGCCGATTTTGAAAGATCCAGTAAGTTCAGTTTCTCTATG
AACTAATCATTTCAAGTCAAAGGCACACTGATGCAAAATCAGTATATGGGACCCCTGTGCTGGATTAGGCAAGGTTTTTCTTGG
AAGCCTTTAAACCAACTCCTTCCNTAAAGGTTATTAAAAGGGCTTATGGGGAGTTATTATTTT

145

MGLPRGSFFVLLLLTAACSGLLFALYFSAVQRYPGPAAGARDTTSFEAFFQSKASNSWTGKGQACRHLHLAIQRHPHFRGLF
15 NLSIPVLLWGLDFTPALWDRLSQHKAPYGRGLSHQVIASLTLNGSESAKLFAPPRDTPPKCIRCAVVGNGGILNGSRQGP
IDAHDYVFRNLGAVIKGFERDVGTSTFYFTVNTMKNLSVSYWNLGFTSVPGQDLQYIFIPSDIRDVYMLRSAILGVFVPEG
LDKGRPHAYFGPEASASKFKLLHPDFISYLTFRFLKSLINTHEFDLYMPSTGALMLLTALHTCDQVSAYGFTSNYWKFS
YFERKMKPLIFYANHDLSEALWRDLHKAGILQLYQR

146

GGGACGTCAGCGGACGGGGCGCTCGCGGGCCGGGGCTGTATGGGGCTCCCGCGGGGTCGTCTTCTGGGTGCTGCTCCTGCTC
ACGGCTGCTGCTCGGGGCTCCTCTTTGCCCTGTACTTCTCGGCGGTGCAGCGGTACCCGGGGCCAGCGCCGAGCCAGGGAC
ACCACATCATTTGAAGCATTTCTTTCAATCCAAGGCATCGAATCTTTGGACAGGAAAGGGCCAGGCTGCGGACACCTGCTTCAC
CTGGCCATTGAGCGGCACCCCACTTCCGTGGCCTGTTCAATCTCTCCATTCCAGTGTGCTGTGGGGGGACCTCTTACCCCA
GCGCTCTGGGACCGCTGAGCCAACACAAAGCCCGTATGGCTGGCGGGGGCTCTCTACCAAGTCATCGCCTCCACCTGAGC
25 CTCTGAACGGCTCAGAGAGTGCCAAGCTGTTTGCCCGCCAGGGACACCCCTCCAAAGTGTATCCGGTGTGCGGTGGTGGGC
AACGGAGGCATTCTGAATGGGTCCCGCCAGGGTCCCAACATCGATGCCCATGACTATGTATTGAGACTCAATGGAGCTGTGATC
AAAGGCTTCGAGCGCGATGTGGGCACCAAGACTTCCTTCTATGGTTTCACTGTGAACACGATGAAGAACTCCCTCGTCTCCTAC
TGGAATCTGGGCTTACCTCCGTGCCACAAGGACAGGACCTGCAGTATATCTTCATCCCTCAGACATCCGCGACTATGTGATG
CTGAGATCGGCCATTCTGGGCGTGCCTGTCCCTGAGGGCTAGATAAAGGGACAGGCCGACGCTATTTTGGACCAGAAGCC
30 TCTGCCAGTAAATTCAAGCTGCTACATCCGACTTCATCAGCTACCTGACAGAAAGGTTCTTGAAATCAAAGTTGATTAAACACA
CATTTTGGAGACCTATATATGCTAGTACCGGGCTTCTATGCTGCTGACAGCTTTGCATACCTGTGACCAGGTGAGTCCCTAT
GGATTTCATCACAAGCAACTACTGGAATTTTCCGACCCTATTTTGAACGAAATGAAGCCATTGATATTTTATGCAAAACAC
GATCTGTCCCTGGAAGCTGCCCTGTGGAGGGACCTGCACAAGGCCGGCATCCTTCAGCTGTACCAGCGCTGACCCCAATGCACT
GAGCGCTTTGCTTCTTCAAGAGTTGCGGCCCTGATCCTCTCAAGTGGCCAAAAGCTTTTAACTTTTCAATCTTCACCTTCCC
35 TTGCCAACAGAGGGCACTGGGGTGAATTCAAGATTTTCATCGAGGTCTGTTCAATATAGGACACCCAGCTTGCTCTGGCTCA
TCCAAGAACTCTTCTGTATCTAAACAATACATCTCAATCTTGGCCAAGGGAAATGGACTGCTTTGCTGGATTGGCACTGAGC
AACTTTAGGAAATGTCGGTGGAGTGTTCAGCAAGATCAGACAGCAGTCCAGGTCAAAGGCAAACACACGCTCCAGCCCAAT
CCTCCTGGTGGCACATCCTACCCAGATGCTAAAGTGATTCAAGGACTCCAGGACACCTCTTAAGAGCCTTCTAAGAACATGA
TAGGCTTACTTCTGCTCCATAATAAAGTGGGAGAAAAAGCCAGAATATAACTTAAGACTAGATAACTGCGTACATGATGGACC
40 ATTTTTTTTTTTTTTGGCTGGGTAGAGAAATCATATAAAACGCAAGGCTGTTTAGCATGGAGATGACTCTCAGAACTGGGAGG
GTCTGGCACTTGATGGGGTTAGTTGCTTGGCAGCTGCCTGCCACTGAGGGAAGTCCATTAGAGATGTATCACCACCTTGTC
ACCAACAGGATGATGTACCAACAGGATGATGTACCAGGTAATAAACCTTCATCCTCAC

148

GGGGATAAGTTTTTTTTTTTTCTTTTTTCTTTCTCTAGTTTTTCTTCTGGCTGAGATTTCCGTGCAAGACAGCACCCAAAT
45 AGACTATTTAGAGTTGACATTTGACATTTAATGGGCGCATGGCTCATTTGTAGATTGAGAAGGTGCGTCTCCCTGCTCCA
AGTCTCATCATGACAGCGTCTGACAGCTGGGAGTCTGTGGCTTCTCTACGACAGAGGCTTAAAGCTGGACACAGATGCACGC
CTAGGCTGGGACAGGATGGGACCCATGCCCTCCTTAGAGGACGGCTTCTGTTAGGAAAGGACACGTTGGGGTGCCTTGC
ATAATAGTTCACTGGTCACCGTGCCTTTATGAGTAGTGTTTTGTGCACTTGCCAGGGTTTTCTCTCTGTGTGCGAGGGAGT

GATTTAAGCAATGGTGTCTGGAGTAAGCCTTACAATTTTAATAGACTTTTCTTATCATATCCCTCATTCTTTCCCTGAAATA
AAAAACACACAAGCAAAAAAATGATAGTTTACATCTCTTAGTTCCCTTGCCCAACAAGATATTCTTAGTTCCACTGG
CCAGGATTTTCTACATAGTCAGAACTTACACATTACTAGAGGCACACCCACCAAGGAGTATTGTGTCTACTTTTATCTGTGCA
CCAGCCACAAATACCCACATTGGAAAGACCCATTTGTGATGGGTAAACATCCCTTCTGTCTCCCAACCCCTGTGACTGCCC
5 TGCATGTGTTCATGACCTCCGAAGGCCCTAATTCATGAAGCAGCAAAACCAGCAGATCTCCACCCCTGCCTCAGGACCTCTG
CTGAAGAGGGGGATGAAGTGGGTCTCCAGGGAGGCAGTGGGGGCTTGTGTGGCAGCTGGCTCGGGAGCCGGCTACAGGAGGGC
AGCTCTGCAGTTGGGAGGGCACCCTCCGAGGAGACCAGGCCCTTACACACCCCTTCTACTTATCATCCCTCACACA
CCCTTGTCCAAGGCTTTATGCATCGGATTATTTTCCAAATCGAGAGGACAGTGATAGATGCATTTTCCCGAGGCTGTCTCAG
AAAGGTCGCTAAATGTATACTGTGTGTCAGAAATGCTGAGATCTCCCCCTTTTGGTTTGTGAGCAGTAAAACTCTTTCCA
10 CTGTGACTTATTTTCTCTCTCAGGCAGCCAGCCACCTGGTCCCTTGTGTGACTCTAGCACAGTGGCCAGGATCCAATACGAGT
CCAGGGGTGACCGCAGGATGGTGGGGCAGCGGGCTTCTCCACCTACCCAGCCACCAAGGCCCTGACGCAGTGCCTCTCTGCAC
CTTCAGCACATCCCTGTGCACAGCTGGAAGGGTGCATGGCCCGCTACCTTTGTTCAGATGGGTGGAAACGCTGATGATACCAG
CTCTCCCTGCGGTGCCCTGCCACGGAGCAGGCATTGTGAACCTGGCTGGTGTGTGAGTCCCACGTGGCATGGCTCCAGCCC
AACCACAGTGGAGACTGGAGACAGGGCAATGAGTCTGGTGGGGGCACGTGGACATGCCCCATAGGGGCCCCACCCAGACTTA
15 ACAGGCAAGGTCTTGGGCATTGCGCGACGAGGACTCAATGCTAAAGCAAGCCTGCCTGGCTCTGTGCCAGGGCCCCCTCTCTG
ATTTACACATCCCATTTTACACAGACCCCTTCTCTTAATAAAGGCTGACAGTCTGTGTGGCAGCCAAGAACCCACACCATGA
AGACAGGGAGTGAAGGGCTTTGTGCCAACTCCAGCACAGCTGCGTTCTGGGGTGTGTGAGAGGCATGTTCTGTCTGTGCGC
TGGTGGTCTCGTGAGACAGTTCCGAGGACGGGGAATTCAGGGTGGTGGGGGCTGAGGCTTATATGTGGAACATGATGCAGAG
TTCGCTGCAGACGGATCTGGATATACATATGTATAATTTGTTACGTGTAATTTAAATATATCTGTTTGGCATCTCATGAGA
20 AGATTATATGTAAGGCTCTGAAGGGAGAGGGAGATGTACATTCTGCCAGGCTCTTGGGGACCTTATCCGAGTCATGAAATGTAT
GACTGTGTATCCAGTGGTGAAGAAGCTACACTCCATGTGTATCACGCTTATGACTCCTAATGTATTTTAAGGCAAAAAATG
TCAGCCGACTCCATCTTACCCCTCGATTCTCGAGTCCAGCCTTTCTGTGCCAGTGCTTCACTGAGCCACAACGCTCTCGCCA
TCGGGACCCGGCTGGGCTCGAGTCTCGGGGCACAGTTGCCATGGAGCCCTCTGGGTCAATCTACAAATGTGCTGAGTGCCAG
CTGAAAACCCACAGGAGATGGAGTACCTTGGCCAAGCTTAAAGAGAAGATTTTCTCAGGGTATTTATAGTGTGTCAGCAGG
25 CTCAGGAAGCAGGATGGAAGATGCATTGAGACTGTTAATTTATTAACAAGGCAAAATGATTTGTGTCTTCTGATGACAGACTA
TTAAGTTTGGGACTTATTTTCCCATTTGAGAAGTTATAATATATATTTAAGATGATAAGTTTCTGCTTAAGTTGTGCTTTCA
GCTTCAATGAGTTTAAAGGACACTAAGGCTAATGATACCAATGAGGGTTGGTTTATATCAAACTGAATAGCTGTGCTTTCTC
CAGTAAATATTTTCTTCTACTGAACATGGAGCCATTATTAAGAGTTGTGTGTTTTTATATGTACATTTGTATATTTTTTGC
TTGTTTGTATGTTCTATTTTCTAATAGTTTCTTTTAGTTTCTTAAAGTTGTGATACTAGATTTAGATTCTGATGCTAACTGCA
30 AATCAGGTTGGTCTCTGCTGGGTCTCTCTGCTTTTATTTTACTTTAAGGACAAGTGTAGTTGTGCTCCACCACCTTTCAAAAA
ATGTGAAACTGCCCTGCCCTCCCTTTTGTGTGACAACACTGTGTACATTGACCACCTTCTACCATACTTTATGTTGTAAATCA
AACTCTTTTGTGGTACATTATCTCATGCTTCTGCAAATTCGAATAAATTCATGGCTTCCAAAAAATAAAAAAAAAAAAAA
150

AAGATCTATGTCTTGGGGGCGCCAGGGCAAGCTCCCGGTGACTGCTTTTGAAGCCTTTGATCTGGAGGCCCGTACATGGACC
35 CGGCATCCAAGCCTACCCAGCCGTCGGGCTTTGTGTGGCTGCGCCATGGCTGAAGGCAGCGTCTTTAGCCTGGGTGGCCTGCAG
CAGCCTGGGCCCCACAATTTCTACTCTCGCCACACTTTGTCAACACTGTGGAGATGTTTGACCTGGAGCATGGTGAAGCAGTGG
CTGTTCTGGGCTGTCTCCCGCTCTCTGTGGGATGGAGGGGCATAGTGTGTACATGACTAGATCTGACCTCCCTCTCTGCAG
GGTCTTGGACCAATTTGCCCGCAGCCTGCGCATGAGGGATAAGAGGGCAGACTTTGTGGTTGGGTCCCTTGGGGGCCACATTG
TGGCCATTGGGGGCTTGGTAAGTCTCTATGGGGCTGGGAGAGGAGGGAGTCCCAAGACAGGAAGAGCTAGCCCCAGCATGT
40 GTGTCACCTTCTGCCCCATCTCCAGGCACTCCAGGGCTCAGGGCTTTGTGAGCTCTTTTCCCTATCTATGAAGTAGGCAGGATGT
GGTCTGCCCTGGCCACGACTCAGGGTGTCTCTCAGTGAAGTGGGGCTCTGTCAAGCACCAGCAAGGGCTTACGAGACAGCAGCA
GGACAAATGACCCTGTATGCACAGTCAAGAGAACATATAGAAACATAAATGAATGAACAAATGATGAACATAGGCAATGAGGA
GCAAGGCCCTGATAGCTCATCCTGGGGTGATAACTAGGGCTATTTCTCTGGGCTTAGGGAGCAGAAGGGCTGGAGTCCCTAC
GGCTGTGCTCTCACAGCCTTCTCTCTCTTCTGAGGAAACCAGCCATGTCTTTGGGCTCTGTGGAGAGCTTTAGCCTTG
45 CACGGCGGCGCTGGGAGGCATTGCCCTGCCATGCCACTGCCCGCTGCTCTGCTCTAGTCTGCAGGCTGGGCCCCGGCTGTTTG
TTATTGGGGGTGTGGCCAGGGCCCCAGTCAAGCCGTGGAGGCATGTGTCTGCGTGATGGGGTCTGAAGGCTTGTGGGAGCT
GTCCACTGGAGCAGCTCATTTGCCAGAGGCAGCTATTTCTATGGCTCCTTTGTGCTGAGGACACTCACTGTGGCTCTGTGGGA
TGAGAGAGGCATGGGGGTGAGCACTTGAACACTGCCTTGGGGCTTGGGTAGGGGAGCCTTTGTCTTTAGTGCAGGACACAC
ATATGCTTACACCTACCTTTATCACCATTGTTTCAATGATGCTTAGCTCCATCCTTGGCTTGGGACCTACTAGGCCTTCCA
50 TCCAACTGGGAAATGGGAGAGCAAGCTGGCCCCATGCTCTTCAAGGTCAGTTCTTATCTGGAGTTGACAGGCGCTACCCCA
GTTGCCATTCTGAAAAATCTCAGCTGCCAGGCTGCCTTTAGGGTCCCTGTAGACCAGGAGAGTTGAGAGGGTGGGGACACA

GAGAGAATAGAGAGGATGTGGGAACGCCAGAGGGCCGAGCGCAGGAGTTCAAGTGGAGGAATGCTGGCTTTGAGCCCTCTAC
ACTGCTGGTTGTATGACCTTGGACAAGTCACTTCACCTCTCTGTGCCTCAGCATCCTCATCTATAAATGGGGATCTCTGAAACC
TTCTTACCTTACCTACCTCACAGGGCTGTTGTGAGGACCCAGGAGCTTTGGATGTGGAAGTAAAGTGCTGCTAAAACCT

151

- 5 MVRSLTAVSASWVQAHPPADMGRRKSKRKPPPKKMTGTLETQTFPCFNHEKSCDVKMDRARNITGVIISCTVCLEEFQTPITY
LSEPVDVYSWDIDACEAANQ

152

- GGCACGAGGGGCTCATCCACCTGCAGACATGGGGCGCAGAAAGTCAAACGAAAGCCGCCTCCCAAGAAGAAGATGACAGGCAC
CCTCGAGACCCAGTTACCTGCCCCCTTCTGCAACCAGAGAAATCCTGTGATGTGAAAATGGACCGTGCCCGCAACACCGGAGT
10 CATCTCTTGTACCGTGTGCCTAGAGGAATTCAGACGCCATAACGTATCTGTGAGAACCCGTGGATGTGTACAGTGATTTGGAT
AGACGCCTGCGAGGCGGCCAATCAGTAGCGACACAGAGGACCCGCCCTGAGCAGCCCGGTACTGTGGATCCAGCTGTTTCG
GTTCTGGTCCAGAGACATTCCAGGGGTCCAGGGTGTGGGTCTGGGCTGTACAGCCGTGTGTGTGTGTGTGTGTGTGTGTGTGT
TGT
TGGCCGTGAGT
15 AAGCCTTTGCTGTGTTCCCTGCAGCCCCCTGGCCCCCAGCCTTGGGGCTCTGGCTCCCCCGGGCGGAATTGGGGGACTGTTTCC
TGACATCCTGGACAAGGGAAGCCCACTAGAGGCTGGAACAGGACCTCTCCAGCCTCCTCACCAGCACCGTGCCCATCTCAACTG
GACTTCCCGCCCTCCTTCTCCACCTTCTAGTGCCCGTGGCCGGGGATTCAAAGCCCGCTTCCCGAGGTCCTGGGCTGGGCCC
TGACAGGGAGCCGCCCCCTCCCATGGTAACCAGGAAGCCGTTTCATGTTTCAAGTTGCTTTTGTAGAGGAAGCAAGGGCTGGG
ATGGGGACAGCTGTCAATCACAAGCCCTTAAATAAAGCAGCCAGCGCACAAAAA

- 20 153

MPPPSDIVKVAIEWPGANAQLLEIDQKRPLASIIKEVCDGWSLPNPEYYTLRYADGPQLYTEQTRSDIKNGTILQLAISPSRA
ARQLMERTQSSNMETRLDAMKELAKLSADVTFATEFINMDGIIVLTRLVESGTLKLSHYSEMLAFTLTAFLELMDHGIVSWDMV
SITFIKQIAGYVSQPMVDVSILQSLAILESMVLNSQSLYQKIAEITVGLISHLQVSNQEIQTYYAIALINALFLKAPEDKRO
DMANAFQKHLRSIILNHVIRGNRPKITEMAHLVYLQVLTFNLLERMMTKMDPNDQAQORDIIFELRRIAFDAESDPSNAPGS
25 GTEKRKAMYTKDYKMLGFTNHINPAMDFTQTPPGMLALDNMLYLAKVHQDTYIRIVLENSREDKHECPFRSAIELTKMLCEI
LQVGELPNEGRNDYHPMFFTHDRAFEELFGICIQLLNKTWKEMRATAEDFNKVMQVVRQITRALPSKPNSLDQFQSKLRLSLY
SEILRLRQSERMSQDDFQSPPIVELREKIQPEILELIKQRLNRLCEGSSFRKIGNRRRQERFWYCRALNHLKVLHYGLDLDNP
QGEVTFESLQEKIPVADIKAIIVTGKDCPHMKEKSALKQNKVELELAFSILYDPDETILNFIAPNKYEYCIWIDGLSALLGKDMSS
ELTKSDLDLFLSMEMKLRLLDLENIQIPEAPPIPIKEPSSYDFVHYG

- 30 154

- GGAGGGAGGTGTAGAAAGAGGTACATGGAGAACAAGTTTGTCAATCCGTCTGAACTTCAGTTGCCTTACCTGTAAGGCAGCCGT
GTCTGTGTTTTTGTCTCGCAGAATTAGAGCCCATTTGGGAACGATGCCACCACCGTCAGACATTGTCAAAGTGGCCATTGAGTGG
CCAGGTGCTAACGCCAGCTCCTTGAAATCGACCAGAAACGGCCCTGGCATCCATTATCAAGGAAGTTTGTGATGGGTGGTCCG
TTGCCAAACCCAGAGTATTATACCTCCGTTATGCAGATGGTCTCAGCTGTACATCACCGAACAGACTCGCAGTGACATTAAG
35 AATGGGACAATCTTACAACCTGGCTATCTCCCGTCCCGGGCTGCACGCCAGCTGATGGAGAGGACCCAGTCAATCAACATGGAG
ACCCGGCTGGATGCCATGAAGGAGCTGGCCAAGCTCTCTGCCGACGTGACTTTCCGCTACTGAGTTTCATCAACATGGATGGCATC
ATTGTGCTGACAAGGCTCGTGAAAGTGGAAACCAAGCTCTTGTCCAGCTACAGTGAGATGCTGGCATTACCCCTGACTGCCTT
CCTAGAGCTCATGGACCATGGCATTTGTCTCTGGGACATGGTTTCAATCACCTTTATTAAGCAGATTGCAGGGTATGTAGCCA
GCCCATGGTGGACGTGTCAATCCTTCAGAGGTCCCTGGCCATCTGGAGAGCATGGTCTTGAACAGCCAGAGTCTGTACCAGAA
40 GATAGCCGAGGAAATCACCGTGGGACAGCTCATCTCACCTCCAGGTCTCCAACCAGGAGATTACAGCTACGCCATTGCACT
GATTAATGCACTTTTCTGAAGGCTCCTGAGGACAAACGACAGGATATGGCAAATGCATTTGCACAGAAGCATCTCCGGTCTAT
AATCCTGAAATCATGTGATCCGAGGAACCGCCCCATCAAACTGAGATGGCCCATCAGCTATATGCTCTCAAGTCTTAACCTT
TAACCTTCTGGAAGAAAGGATGATGACCAAGATGGACCCCAATGACCAGGCTCAAAGGACATCATATTTGAAGTGAAGGAGAT
TGCAATTTGACGCAGAGTCTGATCCTAGCAATGCCCTTGGGAGTGGGACCGAAAAACGAAAGCCATGTACACAAAGGACTACAA
45 AATGCTGGGATTTACCAACCACATCAATCCAGCCATGGACTTTACCCAGACTCCTCCTGGAATGCTGGCCTTGGACAACATGCT
GTACTTGGCTAAAGTCCACCAGGACCTACATCCGGATTGTCTTGGAGAACAGTAGCCGGGAAGACAAACATGAATGCCCTT
TGGCCGAGTGCCATTGAGCTACCAAAATGCTCTGTGAAATCTGCAGGTTGGGGAACACCAATGAAGGACGAATGACTA
CCACCCGATGTTCTTTACCCATGACCGAGCCTTTGAAGAGCTCTTTGGAATCTGCATCCAGCTGTTGAACAAGACCTGGAAGGA

160

TGGGAGCCCGACGTTTCCGGGAGCGCCGCGTGGTTAGCGTCGGCGGCTTTTGGCATGGCGACTTTTTCTGGCCCGGCTGGGCCA
ATCCTGTGCGTTAATCCGCAGGAAGATGTCGAGTTTCAAAGGAGGTGGCGCAGGTTGCAAGCGCATAACCCAGCGAAAAAA
CAAGAACAACCTACTCCTGGAGTAGTCTATGTGCGCCACCTACCTAACCTACTTGACGAAACCCAGATCTTTTCATATTTCTCC
5 CAGTTTGGCACTGTGACACGGTTCAGGCTGTCCAGAAGTAAAGGACTGGAAATAGCAAAGGCTATGCATTTGTGGAGTTTGAG
TCTGAGGATGTTGCCAAAATAGTTGCTGAAACAATGAACAACCTACCTGTTTGGTGAAAGACTCTTGGAGTGTCAATTTATGCCA
CCTGAAAAAGTACATAAAGAACTCTTTAAAGACTGGAATATTTCCATTTAAGCAGCCATCATATCCATCAGTGAACCGGTATAAT
CGGAATCGGACACTAACACAAAAGCTACGGATGGAGGAGCGATTTAAAAAGAAAGAAAGATTACTCAGGAAGAAATTAGCTAAA
AAAGGAATTGACTATGATTTTCCTTCTTTGATTTTACAGAAAACGGAAAGTATTTCAAAAACTAATCGTCAGACGCTTACAAAA
10 GGCCAGGTTTACGTAAGAAGAAAGAAAAAGTTTCAAGTACTCTTGACACTCCTGAGAAGACTGTGGATAGCCAGGGCCCCACA
CCAGTTTGTACACCAACATTTTGGAGAGGCGAAAATCTCAAGTGGCTGAACGAATGATGATGATAAAGATGATGAAATAGTT
TTCAACAGCCCATATCCTGTGTAAAGAAGAAATACAAGAGACTCAACACCTACACATTCACGGAAGAAAGACGAAGAAGC
AGCAATCAGTGATTTCAATGTATTATATTTCTTTTGAAAAATATAATATTTTATGAGAAAAAAGAAAAAAGAAAAA

162

15 GTAAAAATGACTTGGATTGAAAATATGTGGTAGCCTTTTATTTCTACATTAAGTTCTACCTAGGATATTTCCAAGGACTGCCA
CAAAACCCATATGTGCAGTACTTTACTACTTTGGGAAAGCTGCATCTTTCTACCACATTTTAACATCTAATATATTTAATTTCT
TTGAAGAGGGTTCTGTGTACGTTATTGTAGTTCCAGTTTAATATAGTTCTTTGTATCTCTTAACAGGTTGAAGTTATTGCAAA
ACACTCTGGAAGTAATAATTACATCATAATCATTTATTTTAAACTTAAAGCCTAGAAATTTCTAGAAAGAAATAGGAG
ACATCTCAGAGCAATTTGGTTTTGGTGATATGTTCTCAACAGAAAACCAAGTGTAAATGAATATCATGCCTCAGCACTGTCACT
20 TTTAAACCTGTGAGGATCCCACCGTAAATTTGGAATGGGAGTTCTGAATTTTACGTTTGAATGTAAATATAAACTTCA
GTCAATATCCAGGTTTATTGTGTCTTACTATTTAATAATGAGAGAAGTAATGGCAAGGCTTTACTTTTCAGGAAAGGATAGAAG
TATAGATTAAAGACTGGAAGTTTAAATATATTTAGCCCAAAGGTTACTTTGAATTTGAAGTCTTTGCATTGACTGTTTGTGTTT
GGTTTATTTGTTTAGCTTTACAAGGTACACATAAGTTAGGTTGAGGGGTTGTTAACCTTTCCGTGGTCTGCTTTTCATTCGGTGT
GCTTCTGTCTACAGGTAATGGAAAACATAAGTAGAATAGGTGACCTCTTAGTTTGAACCTTATTTAAGTGTGGGGATGAATTTT
25 TCATCAGAAGTGCTTACAGGGTTACTACCTCAGTTTACAATCTACCTGGTCAATTTTATTTCTATCCAGTTCTAAGAACTGC
CTCCACTGTTTATATATTTACATAATTAAACACATTGAGAATGCAACACTATAAAAGCTGGTCAAAATTTTTCAGAGCCCTTATTC
TGTGTGTTTTTGTGTTTTTCTTTTTTTTTTGGAGACAGAGTCTCGTTCGGTCCCCAGCTGGAGTGCAGTGGCGCGATCTCGG
CTCACTGCAACCTCCGCTCTCTGGGTTACCGGATTTCTCTGCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACACACCACC
ACGCCCGCTAATTTTTGTGCTTTTTTAGTAGAGACGGGTTTCGCTATGTTGGCCAGACTGGTCTTGAATGTGACCTCGT
30 GATCCGCCCGCTCTCAAAGTGTGGGATTTCTGTGTGTTTTGTGACCTCCACTTTAGGTAATCATAGGGAGCACATT
TACAGGATGTTCTAATAACATGAAAACAGGCTAGTTTCAAGCAACAGCAATGTCGGTTGGAAGCAGGCGTCATTTGCCTTGAA
AAAAGCCTTTTGACAACATACAGGCATTTCTTTAAACCAGGCTGAAACATTTTATTTCCGAGACTTAACGTTGTGTTTCTGT
TTCTTAAACCTAGCACCTCTGTGTATTTGAAAATAATGAGACATCTTTCATTGGATTTTGGAAAATTGTTCCCATGGGATTCT
AACCTCACTACCAAATGAGTGAAGCTTGATTAAGAGTTCTTCCATATACTAGCCTCCTTGAAGAAGTGATCAGAAGTGATA
35 AGAAGGACAGAAAGGACTATTTTAAAGTTGACTGAAGGAGAAAAAGCAAAATCTTGTGTTTCAATCCCAATTTAGTTAGAACA
AAGTTAAACCCCGTAATCTTAAAGAGAAAATCTTTGGAGGTTTAAATTAACATTTTATACATTTAAAGTCTTGTAAATGGTG
CTTTAAGTGTCAATGTAGCATGTAAAAGGCTTTGTACAGACAGGTAAAAGTTCCATTTCTGAGTGATGAAATGTAACACTTCTT
CATCTTTAACTTGAAATCAAACTATCAGATTTTATTTTGTATAATTTAAGGAAGGTAAAGTTAGGGGACTAGAAGACTCTAA
ATTGGCTTCTACAGATCAATAATTTAAATGTAAGTGTGGGATTTTATAGTTAAATTTATATTTGTGTATATAACATAATTA
40 TCTGTAAATTGTAATAAATATATTTGCAATTATTAATGTTAAGTGAT

163

MARDPPNRPPTTEGTRGLLSCLPDVERATLTLLLDHLRLVSSFHAYNRMTPQNLAVCFGPVLLPARQAPTRPRARSSGPGLAS
AVDFKHHIEVLHYLLQSWDPRLPRQSPDVAPYLRPKRQPLHLPLADPEVVTTRPRGRGGPESPSPSNRYAGDWSVCGRDFLPCG
RDFLSGPDYDHTVGSDEDEDEVEGPRVTGDFEDDFDAPFNPHLNLKDFDALILDLERELSKQINVCL

45 164

ATGGCCCGGACCCCCCAAACAGAGTTCCCCCACCAGTGGGGCACCCAGGGCTCCTCAGCTGCCAGATGTGGAAAGG
GCCACGCTGACGCTTCTCTGGACCACCTGCGCTCGTCTCCTTCCATGCTTACAACCGCATGACCCACAGAAGTTGGCC
GTGTGCTTCGGGCTGTGCTGTGCGGCACGCCAGGCGCCACAAGGCTCGTGCCCGCAGCTCCGGCCAGGCTTGGCAGT
GCAGTGGACTTCAAGCACCATCGAGGTGCTGCACTACCTGCTGCACTTTGGCCAGATCCCCGCTGCCCGACAATCTCCA

- GATGTCGCGCCTTACTTGGCAGCCAAACGACAGCCACCTCTGCACCTGCCGCTGGCAGACCCCGAAGTGGTGACTCGGCCCCGC
GGTCGAGGAGGCCCCGAAAGCCCCCGAGCAACCGCTACGCCGCGGACTGGAGCGTTTGCGGGCGGGACTTCTGCCCTGTGGG
CGGGATTTCTGTCCGGGCCAGACTACGACCACGTGACGGGCAGTGACAGCGAGGACGAGGAGGTCGGCGAGCCGAGG
GTCACCGGTGACTTCAAGACGACTTCGATGCGCCCTTCAACCCGACCTGAATCTCAAAGACTTCGACGCCCTCATCCTGGAT
5 CTGGAGAGAGAGCTCTCAAGCAAATCAACGTGTGCCTCTGAGCCAGATGACGGGGTGGGACCCCGGTTAGTAAGGACCGGGCG
CCCAGTGGCTAAGGCGGTGCCCTGGTGACCAAGGAGAGCCAGACCTGTTGCTCAGGCCGAGCTCCTGGTTGCCAGCGAGTTACC
ACGGGACCACTCGCGTGTATGGCTGAGACTCATTTCCAGTTTCCAGGGCCCGGTATTTGGACACTAGTTGCCAAGTCTGGGGCC
TGGGGATTTTAGGGACCAGCGGTTGTGACCATCTTTCTGAGCACCAGGGCTTCCCTTTTGTGGCAAAAAGGTAGTTCTCG
CGCTTGCTAGGCTGGCCTCTCTTGCCCTCCCTTGGCCGGGGCAACACCAAGTTACTGTGAGCATCACCTGGTGTGGTGAGTCAC
10 CTCCTAGTGGCCCTCTTGCTGCTGCCAACCAGTATAGCTTTGAGCACTGCACTGTTTCTCCCTCCCTTGGACGACACA
AAGACTAGCATGAGGCACCTTTGTGGGGGGCAGCCCTATCTGGGTTCCAGCATGGACACAGGGGTAGCCTGGGGCTTATAG
AGAAACAGCTGGTTTCCCTTACCTTTCCCGGGGAAGACCCACGATTGGCTCTAGTCAGCAAAATGGAGATAACAGAGACTCTGG
CCTTTCCAATCCCCATCTCTTGCCCCCCCCCTGCCCCCCCCCGAAAAAATTGAGCACTTAAACCCCTCCCTTTTGGAGGGG
GCCCCCTGAAGCGTCAGGCTGGGGGCGAGTCTGGTACGGAACATAATTATTGCTCCATGCATGTGTGTGTGTGTGTGTGAGGAC
15 TGGTGTGCGTGGACACGTCTGAAGCAGGCGTGTGGGGCTCTTTCAGGGACCACAGAGGAGGGAGCAGTTTGCAGTGGCCAGCCA
CCCTGAAATCCCCAATAATGGTGCCTCAGTGGGGCCAGAGTTCCAGTGGGAGAGTACGGTTCCCTCCTGTCTCCCTCTCTTT
TCCGCACCTCCATCTTTGTGGATAATAAATAAATATGCACAGGT
166
- TTTTAGGATGGGGTGTGGGAATGAGGCTGTGTGTAGGGAGGGGGTGCGGGGTGGGGACGGAGGGGAGCGTCTTAAGGTCGATT
20 TAGTGTCTATGCTCTTTTACCACCACCACCACCAGAGATGACAGCAAGGATCGGCTAAATACCGCGTGTCTCATCTAGAA
GTGGGAACCTTACAGATGACAGTTCTTGCATGGGCAGAACGAGGGGACCGGGGACCGGGAAGTCTGCTTGAGGGAGGAGGGGTG
GAAGGAGAGACAGCTTCAGGAAGAAAAACAAACACGAATACTGTGCGACACAGCACTGACTACCCGGGTGATGAAATCATCTGC
ACACTGAACACCCCGTCACAAGTTTACCTATGTCACAATCTTGACATGTATGCTTGAACGACAAATAAAGTTAGGGGGGAG
AAGA
25 168
- TTTTTTTTTGTATTGTATACACAGTGGAAGCTGGTTTTATTTGGGAGACAATGGGAGCTTTTACATTGTGTGACAAAGGAGT
ACGAGATCAGTCTTCTCTTTTAGAAAGATTAGTTTGGCAGTTACTTATTTGTAACAGANTTAGACAGCAATCGGGATGCAGG
GGGAGAAGTCAGGTGACTATTAGTCTGCGAGTAATTCTGGGACAAGAGCAGTGGTAATGGAATTNAAAGGATTAAAGTNTTTA
CCAGGTTTTGGCATAAAT
30 169
- MDFPCLWLGLLLPLVAALDFNYHRQEGMEAFKLTVAQNYSSVTHLSIGSVKGRNLWVLVVGFRFPKEHRIGIPEFKYVANMHG
DETVGRELLHLIDYLVTSBGKDPETNLINSTRIHIMSPNPDGFEAVKPKDCYYSIGRENYNQYDLNRNFPDAFEYNNVSRQ
PETVAVMKWLKTETVLSANLHGGALVASYPFDNGVQATGALYSRSLTPDDDVQYLAHTYASRNPMMKKGDECKNMNFPNGV
TNGYSWYPLQGMQDYNIIWAQCFEITLLESCCKYPREEKLPSFWNNKASLIEYIKQVHLGVKGQVFDQNGNPLPNVIVEVQD
35 RKHICPYRTNKYGEYLLLLPGSYIINVTVPGHDPHITKVIIPEKSONFSALKKDILLPFQGLDSIPVSNPSCPMIPLYRNL
DHSAAATKPSLFLFLVSLHLHIFK
170
- GCATTTCTTCTTCTGCGTATGGGACAGGACCCCTTCTGGAATGGGGGTCTTATGACCTACAATCAAAACAAGAATGGAATTC
CCGTGCCCTCTGGCTAGGGCTGTGCTGCCCTTGGTAGCTGCGCTGGATTTCAACTACCACCGCCAGGAAGGGATGGAAGCGTTT
40 TTGAAGACTGTTGCCCAAAACTACAGTTCTGTCACTCACTTACACAGTATTGGGAAATCTGTGAAAGGTAGAAACCTGTGGGTT
CTTGTGTGGGGCGGTTTCCAAAGGAACACAGAAATGGGATTCCAGAGTTCAAATACGTGGCAAAATATGCATGGAGATGAGACT
GTTGGGCGGGAGCTGCTGCTCCATCTGATTGACTATCTCGTAACCAAGTGTGGCAAGACCTGAAATCAAAATCTGATCAAT
AGTACCCGGATACACATCATGCCCTTCCATGAACCCAGATGGATTGAAGCCGTCAAAAAGCCTGACTGTTACTACAGCATCGGA
AGGGAAAATATAACCAAGTATGACTTGAATCGAAATTTCCCGATGCTTTTGAATATAAATGTCTCAAGGCAGCCTGAAACT
45 GTGGCAGTCATGAAGTGGCTGAAAACAGAGACGTTTGTCTCTCTGCAAACTCCATGGTGGTGCCTCGTGGCCAGTTACCCA
TTTGATAATGGTGTTCAGCAACTGGGGCATTATACTCCCGAAGCTTAACGCTGATGATGATGTTTTTCAATATCTTGCACAT
ACCTATGCTTCAAGAAATCCCAACATGAAGAAAGGAGACGAGTGTAAAACAAAATGAACCTTCTAATGGTGTACAAATGGA
TACTCTTGGTATCCACTCCAAGGTGGAATGAAGATTACAATACATCTGGGCCAGTGTTTTGAATACGTTGGAGCTGTCA

- TGCTGTAAATATCCTCGTGAGGAGAAGCTTCCATCCTTTTGGAAATAATAACAAAGCCTCATTAATTGAATATATAAAGCAGGTG
CACCTAGGTGTAAAGGGTCAAGTTTGTGATCAGAATGGAAATCCATTACCCAAATGTAATTGTGGAAGTCCAAGACAGAAAACAT
ATCTGCCCTATAGAACCAACAAATATGGAGAGTATTATCTCCTTCTCTTGCCCTGGGTCTTATATTATAAATGTACAGTCCCT
GGACATGATCCACACATCACAAAGGTGATTATCCGGAGAAAATCCCAGAACTTCAGTGCCTTAAAAAGGATATTCTACTTCCA
5 TTCCAAGGGCAATTGGATTCTATCCCAGTATCAAATCCTTCATGCCCAATGATTCTCTATACAGAAATTTGCCAGACCACTCA
GCTGCAACAAAGCCTAGTTTGTCTTATTTTTAGTGAGTCTTTTGCACATATTCTTCAAATAAAGTAAATGTGAAACTCAACC
CACATCACCACCTGGAATCAGGGATTGCTCACTCCAGGTACTGCAACCCCTAACTCACTCTAGTGGGACCTTGACTGGAGAAAC
TCCACGATCTTCTGAAGAAGAGAAAATGGATGTTTCCAAATCCACAATAAGCAATATGTGGTGATAATGAAAAGAATGATTCA
GTCTTGACGGTGAATGGAAGACACTTACCTAACAAAGTACTGCTCATTACACTCAAATTAATCTTGAAGTAGTCTTAAATGTG
10 TAAGAAGTTAAACTTGAGAAGCAAAAAATGCCGCAAAAAGAAGATCATTTTGTATACAGAGAACCGGATGAATATAAGCAAT
GAAGATGAACATTTATTGATCTTCTACATACAAGACTTCACCATAAAGGCCAGGAGCAGTGCTCAGCCCTTGTAAATCCCAGCACT
TTGGGAGGCCAAGGTGGGCGGATCACCTTGAGGTGAGGTTCAAGACCAGCCTGACCAACATGGTGAAACCTGTCTCTACTA
AATATTAGCGGGGTGTGGTGGCGGCACCTGTAGTCGACGCTTTCGGGAGGCTGAGACAGGAGAATCGCTTGAACCTTAGAGG
CGGAGTTTGCAGTGAGCCGAGATAGTGCCATTGTACTCCAGCTTGGGCAACAGAGTAAGACTCTGTCTC
- 15 172
AGGTTTGAGGGGGGATGCTGTCATGAAAAGGAACCTTTCGCTTTTCAGGTTATGTATGTATAGTTTCTGAACTGGTATCACAAC
CACATGCTGCTTCTGTGTAACTGATTGGTCAAAAAGCCCTGTTCCTCTGAGAATCACCAAAATGCTTCCCAAAGATGCTA
TATCCTTACACCAAACTTAAATCACTGGCTACTGAGAATATGCAGTCTTTGCGCACAAATGCGTCTTCTGTCTGCCCTCCCG
CAGGTCCCAGACACAATTTCTTCATGGCTTCTTCAGCCTTTCACCCCTCAGCCTCTGTCTCAAATTTGAACCCATTTTACATG
20 ACCACACAGGGACACATCATCTCTTTATGGAAGAAGGTGAGATAGAAAACAAATAATGAGGTGTTGATTGATTGAATTTCTCTG
AATCCTAGACCTTAGGATCTTGAATGGCCCCCTTTCTCTGTACGATAGCCCATCCAATGCCTTCCGGCTCCACCGGCTCCTCT
GGAGAGCTCAAATCATGAATAACN
- 174
TAATAAAGTGCTGAATATCTGTCACTTTATTGAATTGAATACTGTACATTACATCAAAATGCAACTGTTTCACATCATCATAAA
25 GTCAATATGTATGACGTTGCACCATCAGAAGTTGGGGACCATCTGTATAAGTTATCACCTTAGTTATAGACTACGTAGGTGT
AGCATGAAAGAGTTAAACCTCCCTCCCCACACATACACAGGGCACAGGGCTGCTGCGCTGGGTGGGAGGGTGCACCT
- 176
ANNANAGTTTGAGGGAAGAGGTAAGAATTGACACAAATAAGTAATGTATGTGCTTAAATGACTGAAGATTGAGTAAAGA
CCAAATGGAAGAAAATCTCATTTTTTTAATTCCNCTCATTTTGATGTTTGTGAAAGTCTGTATTGGTGCAGGCACTTTGGGTG
30 TTTCTGTGGCCGAAGAGGCTCCTGATCTTGCCCTCAGATGAGTTTTTTTTTTGAAAAAGGAGCACAATGGAGAGCAAGAAA
AATAAAGGGAATACAGCCTTTGTGAATGGGAAGGAAGTGTGGTTTCAGTTGCAGAGGTGGGCCAAATAGTTTTGGCCTTTTTT
TTGTCAGGCAGACACATGCACACGTATGTTTATTGCGGTACTATTACAAATAGCAAAGACTT
- 178
AGGATGGGTGTGACCTTAATCTACTGACAAAGACAAAAGAAAATTTCAATTATATGTTGATAGTCCACCAGAGGGGCTTGG
35 AGGGAAACAGTATACCCCCAGTAATGCAAAAGAGGTATAATGGTCCTAGAAGTGATAGGGTTGTGGACAACCTGCAATTATTCA
CATTAACCTGTTCAAATGTAAAGATCATAAGGAGTGGGCACCCACTATTGAAAAGAGGTGGGGCATCCTTGTCTAGGGCACAAGA
ACCAACATAAATGCAATTATGCATGTCAACAAGCCAAGAAAAGGGGACTTAGGGGGTCAGCAACCCTGGTATACTGCAAAATGG
AGGGTGAGAAGGTAAGAACTTGTTTAATTAAGGATGAACCTGTGGGTGATTACNTACCTTTAATGTCCCAGGGGACNTTTGT
TGTTG
- 40 180
TTTTTTTTTGGTTTGAGGGGAATGCTGGAGAAAACATCAACATTGAGTTGACATTTGTTTTGTGTAAGTATAGCTACCATCCA
CTATCATGAATTTTTGTTCATTACAAATGATAGAAAAGCCAGATTCTCAAAATAAAGTATAATTTCTTTGTATTAAATAAATGT
TTATAAATGTTTATGAAGCTCATTACATTATCTTTTTTAAAAAGTAAANTTTTGAACATATGACGCTTTTCATAATTAATG
CTTTTGATATAGATTTGAGGAAAGAGTCTCCAGTTGCACTTCTGAAAATGTAATTTTCAATTAATTTTTCATTGGCATTGAG
45 GGTGACACCTTCCNCTATAATTTCAATTACCCGATTGTAAGGTTTCCCCC
- 181

- MGQQPGKVLGDQRRPSLPALHFIKAGGKKESSRHGGPHCNVFEQALQRPVASFEPQGLSEARWNSKENLLAGPSENDPNLF
VALYDFVASGDNTLSITKGEKLRVLGYNHNGEWCEAQTKNGQGWVPSNYITPVNSLEKHSWYHGPVSRNAAEYPLSSGINGSFL
VRESESSPSQRSISLRYEGRVYHYRINTASDGKLYVSSSRFNTLAEVLVHHSTVADGLITTLHYPAKRNKPTVYGVSPNYDK
WEMERTDITMKHKLGGGQYGEVYEGVWKYSLTVAVKTLKEDTMEVEEFLKEAAVMKEIKHPNLVQLGVCRTREPPFYIITEFM
5 TYGNLLDYLRECNQEVNAVVLlyMATQISSAMEYLEKNFIHRDLAARNCLVGENHLVKVADFGLSRLMTGDTYTAHAGAKFP
IKWTAPESLAYNKFISIKSDVWAFGVLLWEIATYGMSPYPGIDRSQVYELLEKDYRMKRPEGCEPKVYELMRACWQWNPSPDRPSF
AEIHQAFETMFQESSISDEVEKELGKQGVRAVTTLLQAPELPTKTRTSRRAAEHRDITDVPMPHSGKQGESDPLDHEPAVSP
LLPRKERGPPEGGLNEDERLLPKDKKTNLFSALIKKKKTAPTTPKRSSSFREMDGQPERRGAGEEEGRDISNGALAFPLDPA
DPAKSPKPSNGAGVPNGALRESGGSGFRSPHLWKKSSLTSSRLATGEEEGGSSSKRFLRSCSVSCVPHGAKDTEWRSVTLPR
10 DLQSTGRQFDSSTFGGHKSEKPALPRKQAGENRSQVTRGTVTTPPRLVKKNEEADEVFKDIMESSPGSSPNLTPKPLRRQV
TVAPASGLPHKEEAWKGSALGTPAAAEVPTPTSKAGSAPRGTSKQPAEESRVRHKKHSSSESPGRDKGLSKLKPAPPPPPAAS
AGKAGGKPSQRPGQEAAGEAVLGAKTATSLVDVNSDAAKPSQPAEGLKKPVLPAATPKPHPAKPSGTPISPAVPLSTLPSAS
SALAGDQPSSTAFIPLISTRVSLRKTQPPERASGATKGVLDSTALCLAISGNSEQMASHSAVLEAGKNLYTFCVSYVDSI
QQMRNKFAFREAINKLENNLRELQICPASAGSGPAATQDFSLLSSVKEISDIVQR
- 15 182
- ATGGGGCAGCAGCCTGGAAAAGTACTTGGGGACCAAGAAGGCCAAGCTTGCCTGCCCTGCATTTTATCAAAGGAGCAGGGAAG
AAGGAATCATCGAGGCATGGGGGTCCACACTGCAATGTTTTGTGGAAACAAGCCCTTCAGCGGCCAGTAGCATCTGACTTTGAG
CCTCAGGCTCTGAGTGAAGCCGCTCGTTGGAACCTCAAGGAAACCTTCTCGCTGGACCCAGTGAAAATGACCCCAACCTTTTC
GTTGCACGTGATGATTTTGTGGCCAGTGGAGATAACACTCTAAGCATAACTAAAGGTGAAAAGCTCCGGGTCTTAGGCTATAAT
20 .CACAATGGGGAATGGTGTAAGCCAAACCAAAAATGGCCAAGGCTGGGTCCCAAGCAACTACATCACGCCAGTCAACAGTCTG
GAGAAACACTCCTGGTACCATGGGCCTGTGTCCCGCAATGCCGCTGAGTATCCGCTGAGCAGCGGGATCAATGGCAGCTTCTTG
GTGCGTGAGAGTGAGAGCAGTCTAGCCAGAGGTCCATCTCGCTGAGATACGAAGGGAGGGTGATACCATACAGGATCAACACT
GCTTCTGATGGCAAGCTCTACGTCTCTCCGAGAGCCGCTTCAACACCTGGCCGAGTTGGTTTCATCATTCACCGGTGGCC
GACGGGCTCATCACCACGCTCCATTATCCAGCCCCAAGCGCAACAAGCCCACTGTCTATGGTGTGTCCCCAACTACGACAAG
25 TGGGAGATGGAACGCACGGACATCACCATGAAGCACAAGCTGGGCGGGGGCCAGTACGGGGAGGTGTACGAGGGCGTGTGGAAG
AAATACAGCCTGACGGTGGCCGTGAAGACCTTGAAGGAGGACACCATGGAGGTGGAAGAGTTCTTGAAGAAGCTGACAGTCATG
AAAGAGATCAAAACCCCTAACCTAGTGCAGCTCCTTGGGGTCTGCACCCGGGAGCCCCGTTCTATATCATCATGAGTTCATG
ACCTACGGGAACCTCTCGACTACCTGAGGGAGTGCAACCGGAGGAGGTGAACGCCGTGGTGTCTGATACATGGCCACTCAG
ATCTCGTCAGCCATGGAGTACCTAGAGAAGAAAACTTCATCCACAGAGATCTTGCTGCCCGAAAAGTGCCTGGTAGGGGAGAAC
30 CACTTGGTGAAGGTAGCTGATTTTGGCCTGAGCAGGTGATGACAGGGGACACCTACACAGCCCATGCTGGAGCCAAGTTCCCC
ATCAAATGGACTGCACCCGAGAGCCTGGCCTACAACAAGTTCTCCATCAAGTCCGACGCTTGGGCATTTGGAGTATTGCTTTGG
GAAATTGCTACCTATGGCATGTCCCTTACCCGGGAATTGACCGTTCCAGGTGTATGAGCTGCTAGAGAAGGACTACCGCATG
AAGCGCCCAAGAGGCTGCCAGAGAAGGTCTATGAACATCATGCGAGCATGTTGGCAGTGAATCCCTCTGACCGGCCCTCTTT
GCTGAAATCCACCAAGCCTTTGAAACAATGTTCCAGGAATCCAGTATCTCAGACGAAGTGAAGAAAGGAGCTGGGGAACAAGGC
35 GTCCGTGGGGCTGTGACTACCTTGCTGCAGGCCCCAGAGCTGCCACCAAGACGAGGACCTCCAGGAGAGCTGCAGAGCACAGA
GACACCACTGACGTGCCTGAGATGCCTCACTCCAAGGGCCAGGAGAGAGCGATCCTCTGGACCATGAGCCTGCCGTGTCTCCA
TTGCTCCCTCGAAAAGAGCGAGGTCCCCCGGAGGGCGGCTGAATGAAGATGAGCGCCTTCTCCCCAAAGACAAAAGACCAAC
TTGTTACAGCGCCTTGATCAAGAAGAAGAAGACAGCCCCAACCCCTCCCAACGCAGCAGCTCCTTCCGGGAGATGGACGGC
CAGCCGGAGCGCAGAGGGGCCGGCGAGGAAGAGGGCCGAGACATCAGCAACGGGGCACTGGCTTTACCCCCCTTGGACACAGCT
40 GACCCAGCCAAGTCCCAAGCCAGCAATGGGGCTGGGGTCCCAATGGAGCCCTCCGGGAGTCCGGGGGCTCAGGCTTCCGG
TCTCCCCACCTGTGGAAGAAGTCCAGCACGCTGACCAGCAGCCGCTAGCCACCGGCGAGGAGGAGGGCGGTGGCAGCTCCAGC
AAGCGCTTCTCGCTCTTGCTCCGTCTCTGCTTCCCATGGGGCCAAGGACACGGAGTGGAGGTGAGTCACGCTGCCTCGG
GACTTGCAGTCCACGGGAAGACAGTTTGACTCGTCCACATTTGGAGGGCACAAAAGTGAAGAAGCCGCTCTGCTCGGAAGAGG
GCAGGGGAGAACAGGTCTGACAGGTGACCCGAGGCACAGTAACGCCTCCCCCAGGCTGGTGAAGAAAGAAAGTGAAGAAAGTCT
45 GATGAGGTCTTCAAAGACATCATGAGTCCAGCCGGGCTCAGCCCGCCCAACCTGACTCCAAAACCCCTCCGGCGGAGGTG
ACCGTGGGCCCTGCCTCGGGCCTCCCCACAAGGAAGAAGCCTGGAAGGACAGTGCCTTAGGGACCCCTGCTGCAGCTGAGCCA
GTGACCCCCACAGCAAAGCAGGCTCAGGTGCACCAAGGGGACACAGCAAGGGCCCCGCGAGGAGTCCAGAGTGAGGAGGCAC
AAGCACTCTCTGAGTCGCCAGGGAGGACAAGGGGAAATGTCCAAGCTCAAACCTGCCCGCGGCCCCACAGCAGCCTCT
GCAGGGAAGGCTGGAGGAAAGCCCTCGCAGAGGCCCGGCCAGGAGGCTGCCGGGAGGAGTCTTGGGCGCAAAGACAAAAGCC
50 ACGAGTCTGGTTGATGCTGTGAACAGTGACGCTGCCAAGCCAGCCAGCCGGCAGAGGGCTCAAAAAGCCCGTCTCCCGGCC

ACTCCAAAGCCACACCCCGCCAAGCCGTCGGGGACCCCATCAGCCCAGCCCCGTTCCCTTTCCACGTTGCCATCAGCATCC
TCGGCCTTGGCAGGGGACCAGCCGTCTTCCACTGCCTTCATCCCTCTCATATCAACCCGAGTGTCTCTTCGGAAAACCCGCCAG
CCTCCAGAGCGGGCCAGCGGCCATCACCAAGGCGTGGTCTTGGACAGCACCAGGCGCTGTGCCTCGCCATCTCTGGGAAC
TCCGAGCAGATGGCCAGCCACAGCGCAGTGTGGAGGCCGGCAAAAACCTCTACACGTTCTGCGTGAGCTATGTGGATTCCATC
5 CAGCAAAATGAGGAACAAGTTTGCCTTCCGAGAGGCCATCAACAACTGGAGAATAATCTCCGGGAGCTTCAGATCTGCCCGGCG
TCAGCAGGCAGTGGTCCGGCGGCCACTCAGGACTTCAGCAAGCTCCTCAGTTCCGTTGAAGGAAATCAGTGACATAGTGCAGAGG
TAGCAGCAGTCAGGGGTACAGGTGTACAGGCCGTCGGAGCTGCCTGCAGCACATGCGGGCTCGCCCATACCCATGACAGTGGCTG
AGAAGGGACTAGTGAGTCAGCACCTTGGCCCAGGAGCTCTGCGCCAGGCAGAGCTGAGGGCCCTGTGGAGTCCAGCTCTACTAC
CTACGTTTGCACCGCCTGCCCTCCCGCACCTTCCCTCCTCCCGCTCCGCTCTGTCTCGAATTTTATCTGTGGAGTTCCTGCT
10 CCGTGGACTGCAGTCGGCATGCCAGGACCCGCCAGCCCCGCTCCACCTAGTGCCCCAGACTGAGCTCTCCAGGCCAGGTGGGA
ACGGCTGATGTGGACTGTCTTTTTCATTTTCTCTCTGGAGCCCCCTCTCCCCGGCTGGGCCCTCTTCTTCCACTTCTCCA
AGAATGGAAGCCCTGAACAGGCTGTGTGTACAGGCCCTCTGCCTGCACCTCCCTGGCCCTGCGCTGTGTGTGAAGACATG
TTTCAAGAAGCCGCAATTTCCGGAAGGGCATGCACGGGCCATGCACAGGCTGGTCACTCTGCCCTCTGTGTGTGTGTGTGTGTG
GGTGCACCTCGCCATTTCTTCAGTGCAGGACAGCTCTTGATTTGGGTGGAAAACAGGGTGTAAAGCCCAACAGCCTTTGGGTG
15 CTGGGCAGGTGGGAGCTGAAAAGGATCGAGGCATGGGGCATGTCTTTCCATCTGTCCACATCCCCAGAGCCAGCTCTTGCTC
TCTTGTGACGTGCACTGTGAATCCTGGCAAGAAAGCTTGAGTCTCAAGGGTGGCAGGTCACTGTCACTGCGGACATCCCTCCCC
CAGCAGAATGGAGGCAGGGGACAAGGGAGGCAGTGGCTAGTGGGGTGAACAGCTGGTGCCAAATAGCCCCAGACTGGGGCCAGG
CAGGTCTGCAAGGGCCAGAGTGAACCGTCTTTCACACATCTGGGTGCCCTGAAGGGCCCTTCCCTCCCCACTCTCTAAG
ACAAAGTAGATTCTTACAAGGCCCTTTCCCTTTGGAACAAGACAGCCTTCACTTTTCTGAGTTCTTGAAGCATTTCAAAGCCCTG
20 CCTCTGTGTAGCCGCCCTGAGAGAGAATAGAGCTGCCACTGGGCACCTCGCGACAGGTGGGAGGAAAGGGCTGCGCAGTCTCTG
GTCTGGCTGCACTCTTGAATGGGCGAATGTCTTATTTAATTACCGTGAGTGACATAGCCTCATGTTCTGTGGGGGTGATCAG
GGAGGGTTAGGAAAACCAAAACGAGCCCTGAAAGCCTCACGTATTTACAGAGCACGCTGCCATCTTCTCCCCGAGGCTG
CCCCAGGCCGAGCCAGATACCGCGGGCTGTGACTCTGGGCAGGACCCGGGGTCTCTGGACCTTGACAGAGCAGCTAACT
CCGAGAGCAGTGGGCAGGTGGCCGCCCTGAGGCTTCACGCCGAGAAGCCACCTTCCCGCCCCCTCATACCGCCTCGTGCCAG
25 CAGCCTCGCACAGGCCCTAGCTTTACGCTCATCACCTAAACTTGTACTTTATTTTCTGTAGATAATGGTTTCCCTCTGGATCGT
TTTATGCGGTTCTTACAGCACATCACCTCTTTCCCCCGACGGCTGTGACGCAGCGGAGAGGCACAGTACCGACAGCGGCCCT
TGAAGACAGAGCAAAGCCCCACCCAGGTCCCCGACTGCTGTCTCCATGAGGTACTGGTCCCTTCTTTTGTAAAGTGATG
TGCCACTATATTTTACAGTATCTCTTGGTATGCATCTTTATAGACGCTCTTTTCTAAGTGGCGTGTGCATAGCGTCTGCCC
TGCCCTCGGGGGCTGTGGTGGCTCCCCCTCTGCTTCTCGGGGTCCAGTGCAATTTGTCTGTATATGATTCTCTGTGGTCTT
30 TTTTGAATCCAATCTGTCTCTGTAGTATTTTTTAAATAAATCAGTGTTTACATTAG

184

GTAGAAAAAAGGTCTCACTGTGCTCAGGCTGGTCTTGAGCTCCTGTCTGGGCAACTTGGCTAAACCTCATCTGTACAAAAAAT
ACGAAAACCTAGCTGGGTGTGGTGTGTCATCCGTGGTCCCAGCTACTAGAGAGGCTGAGGTGGGAGGATTCTTCCACCCAGG
AGGTCAAGGCTGCAGTGAGCCATGATCTCATCACTGTACTCCAGCTTGGGCAACAGAGTGAGACCTGTCTCCAACAACAACA
35 AAAAAAACAACCAAGAAATCTAATCTATCTAGGCAACTTCCAGACCTTAGGTTTGATCCCCACTTTGTCACTTCCCTACAT
GTATGATGTTGGATCTCAATTTCAAACAGTGACATGAGTACCATAACCTTTCAAAAAGTATCTATGGTGTAGCCTACTGGGGT
GCTGGCTATTGTTCTNAGTGCTTTGGGACAGATTAAGTAAACCAAGCAAAACCAAGATCTCTGGCCCTGTGGGAATTCATATT
CTTGGTGAGGGCCCATTTACTTTTCTCTTTTGGATTATATATCCCTTTTGGTTATAATTACG

186

AACTTATAACATTTTGGGTTTATATTTAATATAGTACTTCTCATCAGGAGGATGTTACTCAGTTAATAATAAGTTTCTTCAA
AACATTAAATCTCTTTTCCATGTCAATGTCTATAGTTTCTTTTAAACATTAAATTTTTCGCAATTCAGTATTAGATACAC
TGAATACATTTNTCTAAATGTTTTCNCCCTAGNGGATAAAAGTTTCTCTTTTACCTGGACTCTCTGGATATCTAAATATAATA
TTAACTNNGGGGGGACAACAAAACAAAATCTAAATTTAAAGGNCCTCCCTCAACGGGTTTCGGGGGGCCGGGTGGATTATAACC
187

45 MEPQEERETQVAWLKKIFGDHPIPIQYEVNPRRTTEILHLSERNRVRDRDVLVIEDLKQKASEYESEAKYLQDLLMESVNFSP
ANLSSTGSRYLNALVDSVALETKDTSLSASFIPAVNDLTSDFRTKSKSEEIKIELEKLEKNLTATLVLEKCLQEDVKKABLHL
STERAKVDNRQNMDFLKAKSEEFRTGKAAEEQLSARGMDASLSHQSLVALSEKLARLKQQTIPKKKLESYLDLMPNPSLAQ
VKIEEAKRELDSEIHELTRRVDMEL

188

ACTGGGTNGGAATGGAGCAGGGGCTAGAGGAATCTGGAAGCTAATTCCTTTTAGGCTTTTCCAACAGAGGGAAGCAGAGGGACT
TCTGGGTCTTGAAGTTTGGAGCTTACGGAGGCGA

201

5 MNGKRPAPGPARVGKKGKKEVMAEFSDAVTEETLKKQVAEAWSRRTPPFSHEVIVMDMPFLHCVIPNFIQSQDFLEGLQKELM
NLDPFHEYNDLYKFQQSDDLKKRREPHVSTLRKILFEDFRSWLSDISKIDLESTIDMSCAKYEFTDALLCHDDELEGRRIAFIL
YLVPSWDRSMGGTLDLYSIDEHFQPKQIVKSLIPSWNKLVEFVSPVFSHQVSEVLSEEKSRLSISGWFHGPSLTRPPNYFEPP
IPRSPHIPQDHEILYDWINPTYLDMQVQIQEEFESESEILLKEFLKPEKFTKVCEALEHGHVWSSRGPPNKRIFYEKAESK
LPEILKECMKLFRESEALFLLLSNFTGLKLHFLAPSEEDENDKKEAETDITEEGTSHSPPEPENNQMAISNNSQOSNEQTDPE
PEENETKKESSVPMCGELRHWTGHYTLIHDHSAEFALDLILYCGCEGWEPEYGGFTSYIAKGEDEBELLTVNPESNSLALVY
10 RDRETLKFVKHINHRSLQKKTFFNRTGFWDFFSIYYE

202

CACGATAAAGGGGACATGCCGGGAGTTGCAGTACCCTCAGGAAGAAGTCATTGTTCATGGACATGGACCCTTTTCTTCACTGTGT
GATCCCAAACTTTCATCCAAAGCCAAGACTTCTTAGAAGGGCTTCAGAAGGAAGTCATGAAGTTGGACTTCCATGAGAATCTGAT
GATTTGAAGAAGAGAAGAGAGCCTCACATCTCCACTTTAAGGAAAAATCTGTTTGAAGATTTCCGGTCTCGGCTTTCTGATATT
15 TCTAAAATFGACCTGGAATCAACCATTGACATGTCTGTCTAAATATGAATTCATGATGCCCTGTGTGCCATGATGATGAG
CTGGAAGGGCGCCGGATTGCCCTTCATCTGTACCTGGTTCTTCTCGGACAGGAGCATGGGTGGTACCCTGGACCTGTACAGC
ATAGATGAACACTTTTCAGCCGAAGCAGATTGTCAAGTCTCTTATCCCTTCGTGGAACAAACTGGTTTCTTTGAAGTATCTCCT
GTGTCTTTTACCAGGTGTCTGAAGTGTCTGTGAAGAAAGTCACGTTTGTCTATAAGTGGCTGGTTTTCATGGTCCATCATTG
ACTCGGCCCTCCCAACTACTTTGAACCCCCATACCTCGGAGCCCTCACATCCCACAAGATCATGAGATTTGTATGATTGGATC
20 AACCTTACTTATCTGGACATGGATTACCAAGTTCAAATTCAAGAAGAGTTTGAAGAAAGTTCTGAAATTTCTCTGAAGGAGTTT
CTTAAGCCTGAGAAATTCACGAAAGTCTGTGAGGCTTGGAGCATGGACATGTGGAATGGAGCAGCCGAGGTCCCCCTAACAAA
AGGTTTATGAGAAAGCTGAGGAGAGTAAGCTTCTGAGATATTGAAGGAGTGCATGAAGTTATTTTCGCTCTGAGGCATATTCT
TTGTGCTCTCCAACCTTCACAGGCCTGAAGCTTCATTTCTTGGCCCCCTTCGGAAGAAGATGAGATGAATGATAAAAAAGAGGCA
GAAACCACTGATATCACTGAAGAAGGGACTAGCCATAGTCTCTCGAGCCAGAGAATAATCAGATGGCCATCAGCAACAACAGC
25 CAACAGAGCAATGAGCAGACAGACCAGAGCCAGAGGAAAATGAAACAAAGAAAGAAATCAAGTGTTCCTATGTGCCAAGGGGAA
CTGAGGCATTGGAAGACCGGTCACTACACTTTAATTCATGACCATAGCAAGGCTGAATTTGCCCTAGACTTAATTCGTACTGT
GGCTGTGAAGGCTGGGAGCCAGAATATGGCGGTTTACTTCTTACATTTGCCAAAGGTGAAGATGAAGAGCTGCTAACAGTGAAT
CCAGAAAGCAATTTCTTTGGCATTGGTCTACAGAGACAGAGAGACTCTGAAATTTGTCAAGCATATTAACCAACCGAAGCTGGAA
CAAAAGAAAACCTTCCCAACAGAACAGGTTTCTGGGACTTTTCATTCTATTTATGAATGACAGCACTGGGCAAGCTGAAC
30 AAAAATGTGACCTTCTGAATTAAGTGGGAAGTCTGAAAGAGCTAAGCATGGAGTCAAGGAGAACTACATGGTAGCTTGCTGAC
AGTGTCTTTAAACCTGGTTGTCTTTTACTAGGACTCATAATGATTGTCTCAACCGAGACCTTGAGCTTGCAGCTACGTACTTA
TCTCTTGATTAAAAAAGTTGGCTTTTCTTTTAAACATTTAGTCTTTTTCATATTTGGCTTCTTCAGTGAATTT
TTTAAGTTCAATTTGTTTATTGAGGTAAATATTTATAACATAAACTGACCAGCTTACCCATTTTAAATATGCAATTCAG
TGGATTAAAGTACATTCTCATTTGTTGTCCAGCCATCACCATCATCCATCTCCAGAAGTTTCCATCTTCCCAAAATCTGTGCCCCA
35 TTGAACAATAACTCCCACCTCCCTTCCCTTAGCAACAGCCATACTTTTGTCTCTATCATCAACTTCACTACTCATATTTCT
CATGTAAGTGAATCATACAGTATTGTCTCTTTGTGACTGGTTTCACTTAGCATAAAGTCTTTAAGATGCATCCATGTTCCA
GTGTTTCGGTTTCTTTAGAAAACTCATACGTGATTGCAGCCGGGCATGGTGGCTCACGCCTGTAATCCAGCACTTGGGAGGC
CAAGGCAGGCGGATCACCTGAGGTGAGGAGTAAACACCAGCCTGACAAACATGGAGAAACCCATCTCTACTAAAAATACAAA
ATTAGCTGGGCGTGGTGGCACATGCCTGTAATCCCAGCTACTCAGGAGGCTGGGGCAGGAGAATTGGTTGAACCCAGGAGGCGG
40 AGGTTGCAGTGAGCCGAGATTGTGCCACTGCACCTCCAGCCTGGGCAACAAGAGTGAACCTCTGTCTC

204

GAGTAGGAGTGCCCTCCTTGTCTGCACTGCTGGTATGGGGTTAGGCCAGGTAGGACATTCAGAGGGGCTTCTGAAAACCAAGAG
TCCCCTGGGGAAAGGGAAACAGAGTAAGGCAGGCCTTGTCTCACTGCCCTTAAGGGAACCTGGTCACTCGGCACCTTTTAAGCCT
CAGTTTCTCCAGTTCAATAATAAGGACAAGAGCTTTTCCCATGCAATCTCTTTTCCCGGAAAGTTGACTGAGGTGACCAGTAA
45 TAGAATTGAAAAGGGAGAGTGTCTTCAAGTCAATGTGGCATCTGGATTGGGTCTTGAACAAAAACAGGACATTAGTGGGAAA
ATTGGAAATCTGAAAAAGTCTGAATTTAGTTAATAATACCAATTTCACTCTTGGTTTGTACAGATGTACCATGGTGATGTA
AGATGTTGACCTTGGGGTAGGCTGGGTGAAGGGTATACAGGAACCTTTGTACTATCTCTGCAACTTCTCTGTAATCTAGTAT
CATTCCAAAATAAAAGTTTATTTAATTTAAAAA

205

MNKFPRRIPQKSCPRILCWSCQEVSPREVADAICQAI VLSAPGPHAVLLVTQLGRFTDEDQQVVRRLQEVFVGVLGHTILVFTR
KEDLAGSSLEDYVRETNQALAWLDVTLARRHCGFNNRAQGEQEAQLRELMKVEAIMWENEGDYYSNKAYQYTQQNFRLKEL
QERQVSQGGSEDVPGEE SWLEGLS IQKESEE AHRCLLGKADL

206

- 5 GAAAACATTTTGCTGAAAAATATAAGCAAACATCGGCCCTTGCTCCTCTGTGTTCATACACTGTGGAAGCTTTTCTCTGCCTCCT
CCGTGAGAGTGCCTGGCCGGGAGACCAGAAACGTGGTCTTTCTCTTGCTGTGAGCTGGTGACAGATGGAGGAAGAAGAAATA
TGAACAAATTCCCAAGGAGAATCCCCAGAAAGAGCTGTCCAGGATCCTGTGCTGGAGCTGTGAGGAGGTCTCGCCAGAGGTGG
CAGACGCTATCTGCCAAGCCATCGTCTTATCCGCCCCAGGGCCCCACGCCGTGCTCCTGGTGACACAACCTGGGCCGGTTACCGG
ATGAGGATCAGCAGGTGGTCAGGCGCCTGCAGGAGGTCTTTGGAGTGGGGGTCTGGGTACACCATCTGTGTTCACCCGGA
10 AGGAAGACCTGGCTGGCAGCTCCCTGGAAGACTATGTGCGAGAGACCAACAACAGGCCCTTGCTGGCTGGATGTGACCCCTTG
CACGGCGCCATTGCGGCTTCAACAACAGGGCACAGGGGGAGGAGCAGGAGGCCCAACTGCGAGAGCTCATGGAGAAAGTTGAAG
CCATTATGTGGGAAAACGAAGGAGATTATTACAGCAACAAGGCTTACCAATATACCCAGCAAACTTTCCGGTGAAAGAACTAC
AGGAAAGGCAAGTAAGCCAGGGCCAAGGCTCTGAGGACGTGCTGGTGAGGAGTCTTGGCTGGAAGGACTGTCCAGATCCAGA
AGGAATCTGAGGAAGCCACAGATGCCCTGCTGGGGAAGGCTGACCTTTGAGCCTGTGCTGGACTTGAGCCAAGGACACCATCAG
15 CCTTTGACCCCCCTGTGTCCAGCCCTCTGTTCTCTTTCCATCCCATGGAGTGTCTCCAGTCTCCAGGTATGACGTCTGGT
GTAGGAAGAGGGGCATGGGGCTTGAGGACAGGCTCCAGCATGCCAGATCACATTCCAGTTCCCTTCTCTCCAGAACCA
TCTGGGTCTCTCTCTGTGCCCCAACAGCAGCGTGTCTTCTTATTTGACCTTCTATCTCATTGGCATTGCCCTTTTCATGAGT
TACTTGTCTATACTCTGATAGAACCAAAGTTTCTTGAGGGCAGGCCGTGTATCTTTTACTTCTGTGTAACCTCCACAGCTTTTGA
CTGCATTGGAGTCATGCAGTCACGTTTGGTGCGGCAACTCCATGCAAGCTGTGTGACCAGATGACTTCTCATCTCTCCCCACACC
20 ATGATGGTTTAAAGGTGCCTGTGTGTCTGGTGAGACAGTGTGAGGAGATGATCAGAAATCTCAGGGCACAGCCTCAGGTCTAC
CTGGACACTCATCCACTGTGGAGATGCTCAACAGCTAGACACAGATGCTCACAGCCGACCTGTGGCCTCTTGCTCTGAAGCCA
TCCCCATCTTCCCTAGACACAGCAGACATCTGAGAAAGCTTCAGCATTCTCTTGCTAACAGATTGAGAAAGTGTCTCAAAG
CAGAGCAGAGTTATTGTGTGTGTGCTGAAGACAGCCTTTGTGCCACAATCACTTATTAATAAGCGATCAATTTCCCATTTGA
ACTGAACATGCAACATTTATCATACATTCAGCTCTCATTACACTCTTAAGATTGGTCAAGATTTTATTTCTGTTCATGTCT
25 TTCTACTTTTCTACTCTGTATAGAATAAAATATTGATTGATTACAGTGGCTTTGACTATAATGTGGGAGCCAATTTTGGCTC
AGTCTTCATTTTATATTTACCTTGTATTCTCAGGCATTTTCTTCTATGTGAGAGTTAAATCATTCGTAAATTTCCCCC
CAAAATCATATTGGTATTCTAGTTGGCAATGTCTTACATTTATGTTAAGTTTGAGGGAATTGGTAGTTCAAGTATAAGTTAATT
AAGGCCATTTTATTTCTAAGTGAACAGACTTGAAACTCCAGAGCTACTGAAGTAAAGTTAGAATCATTTGCATTTTTCATTAG
ATAGGAGATAATTTGTAAATTTTGATGCTATTATTTTAACTCTATTAGCTTAAGTAATGTCATAATAGAAAACACAAGCATTT
30 GACCAATGAGATCCATTACAGCGACTAACTGGCAAGGCACTCAAAACATGTCATTATTAATGTTGTTTTTATTTACCTCAAA
AAAAAAAAAAAAAAAAAAAA

207

MVVIDVKMLSGFTPTMSSIEELENKGQVMKTEVKNDHVLFYLENVFGRADSF TFSVEQSNLVFNIQAPGMVYDYIEKDGAEFL
LTN

35 208

- GGCAGAGGCCCGCTTGGGGTGTGTGCTGCCCCGCTGCGATGGAGGTCTTAAGAGCAAGGGGGGAAGAGGGGCTGGCTCTGG
TAGATTGCAAAATGACCTGCTTTTCTGTATCCCGAGCTGCGTTTCGACCCCTGTGCGATAGTAAATCCAGTAAGGTACCTGC
CGTGGGCAGATCTGAGCTTCTCTTGGACACCCCATACCCAGCTCCTCAGGCCACCCCTAAGTACAATGTTCTCTACCTA
AGAAGGCATCTGGATTCTCTTCTTGGAAATAGTAAAGAACTACTCTTTGACTGTTTTTGACCTCAGTGAACCTCAAAT
40 ACACTGGAATTCGCAATAAATCCAGTATGGTGGTTATAGATGTAATAATGCTATCAGGATTACTCCAACCATGTCTCATCCATTG
AAGAGCTTGA AAAACAAGGGCCAAGTGATGAAGACTGAAGTCAAGAATGACCATGTTCTTTCTACTTGGAAAATGTTTTTGGTC
GAGCAGACAGTTTCACTTTTCTGTGAGCAGAGCAACCTTGTTTCAACATTCAGCCAGCCCCAGGCATGGTCTACGATTACT
ACGAAAAAGATGGAGAAGCATTTCTTTTAAACAACTGATTCTTCTGTATCAAACTGGAAAAAATCATGAACCATCTGACATC
GTGAACAGTCTGCACTGGGCTATGGTTTCTTGACAAGTCTTATTTCTTATCATCCCATTAATGTTGTCTATTTTGCAAAAAA
45 AAAAAAAAAAAAAAAAAAACTCGAGACTAGTTCTCTCTCATAAATTAATTAACCGCTCATTATCGACCTCCCCACCCCA
TCCAACATCTCCGATGATGAACTTCGGCTCACTCTTGGCGCCTGCCTGATCCTCAAATCACCACAGGACTATTCTTAGCC
ATGCACTACTCACCAGACGCTCAACCGCTTTTATCAATCGCCACATCACTCGAGACGTAAATTATGGCTGAATCATCCGC
TACCTTCAGCCAATGGCGCCTCAATATCTTTATCTGCCTCTCTTACACATCGGGCGAGGCCTATATTACGGATCATTTCTC

CTCAATTGATCCAATAACTTGACCAACGGAACAAGTTACCCTAGGGATAACAGCGCAATCCTATTCTAGAGTCCATATCAACAA
TAGGGTTTACGACCTCGATGTTGGATCAGGACATCCCAATGGTGCAGCCGCTATTAAAGGTTGCTTGTTCACGATTAAAGTC
CTACGTGATCTGAGTTCAGACCGGAGTAATCCAGGTCGGTTTCTATCTACTTCAAATTCCTCCCTGTACGAAAGGACAAGAGAA
ATAA

5 214

GTGTTTTGGTGTGGTGTGATTATGTGATTGTTAATGATTTATTCATATATCTTACATTAAGACCTATGGTGTCTTTTAACC
TAATGTGTCTCTTTGAAGTCAAGCAAAACCGTAAGAATGGTTAAAAAATTAACAGCTTTGGTTCAATTCCTTTATAAACACATAA
TATAGGTATTACTGATCGAAGGAAAGTTCAGTGAAGATGCTGGTAGTTTACAGCACTCCGAGAATTAACAATGAGATGTAA
AGGTAACAGCTCTAAAAATTCAGACATTTGCCCTTATTAAGCTCTTTTCTACAAGAGACTTCATTACATTGAAAAATGAAACCTG
GCACCCATCCTAGGAGAAATTATCTGTGCAGAAGTTCTGATTTACAGATTTGTGATCTTAGGGAACAGTGTGCACACCCAGGGTG
AAATTTTC

10

216

GAACACCCCGCTCATCTGAGAAAGGGCAGCATTGCTCCCTCAGCAAAATCAGGGGGAGAGGGAGCTTTCCAGGCTTTTAATCAC
TCTCTGATTTCAGCGCCCAAAACCCAGGGACAGGAAGCCAGCAGATTTTGCTCAGAACTGAGATAATATATTTTCTAACCCAGA
AAACACATCTCAGCTAGTTACCACCTCATTCATCTTTCCAGAGTGGTTATAATTACAGTATTCATTTTCATAATTGCTGCCTATT
TGCAAAATACAAATCCCTGAATCTTAAAATAAATTACAGTTCTCTTTTCTATTTAACAGCCATTGTGC

15

217

MSSHLVEPPPLHNNNNCEENEQSLPPPAGLNSSWVELPMNSSNGNDNGNGKNGGLEHVPSSSSIHNGDMEKILLDAQHESGQ
SSSRGSSSHCDSPSPQEDGQIMFDVEMHTSRDHSSQSEEEVVEGEKEVEALKKSADWVSDWSSRPENIPPKEFHFHRHFKRSVSL
MRKSGAMKKGIFSAEFLKVFIPSLFLSHVLALGLGIYIGKRLSTPSASTY

20

218

GCGGCGGACTCGGCTTGTGTGTGTGCTGCGCTGAGTGCCGGAGACGGTCTGCTGCTGCCGAGTCTGCCAGCTGTCCGACGAT
GTCTGCCACCTAGTCGAGCCGCGCGCCGCTGCACAACAACAACAACACTGCCAGGAAAATGAGCAGTCTCTGCCCGCCGCGC
GGCGCGCTCAACAGTTTCCGTGGGTGAGCTACCCATGAACAGCAGCAATGGCAATGATAATGGCAATGGGAAAAATGGGGGGCT
GGAACACGTACCATCTCATCTCCATCCACAATGGAGACATGGAGAAGATTCTTTTGGATGCACAACATGAATCAGGACAGAG
TAGTTCCAGAGGCAGTTCTCACTGTGACAGCCCTTCGCCACAAGAAGATGGGCAGATCATGTTTGTATGTGGAAATGCACACCAG
CAGGGACCATAGCTCTCAGTCAGAAGAAGAGTTGTAGAAGGAGAGAAGGAAGTCGAGGCTTTGAAGAAAAGTGGCGACTGGGT
ATCAGACTGGTCCAGTAGACCCGAAAACATTCACCCAAGGAGTTCCACTTCAGACACCCCTAAACGTTCTGTGTCTTTAAGCAT
GAGGAAAAGTGGAGCCATGAAGAAAGGGGTATTTTCTCCGAGAAATTTCTGAAGGTGTTTCTCATCTCTCTCTCTCTCTCA
TGTTTTGGCTTTGGGGCTAGGCATCTATATTGGAAGCGACTGAGCACACCCCTCTGCCAGCACCTACTGAGGGAAAGGAAAAGC
CCCTGGAAATGCGTGTGACCTGTGAAGTGGTGTATTGTGCAGTAGCTTATTTGAACCTTGAGACCATGTGAAGCATGACCCAAC
CTACCACCTGTTTTTACATATCCAATTCCAGTAACCCCTCAAATTCATATTTTATTCAAACTCTGTTGAGGCATTTTACTAAC
CTTATACCTTTTTTGGCCTGAAGACATTTTAGAATTTCTTAACAGAGTTTACTGTTGTTTTAGAAATTTGCAAGGGCTTCTTTTC
CGCAAAATGCCACCAGCAGATTATAATTTTGTGCGCAATGCTATTATCTCTAATTAGTGCCACCAGACTAGACCTGTATCATTTCA
TGGTATAAATTTTACTCTTCCAACATAACTACCATCTCTCTCTTAAACAGAGATCAGGTTAGCAAAATGATGTAAAAGAAGCTTT
ATTGTCTAGTTGTTTTTTTCCCCCAAGACAAAGGCAAGTTCCCTAAGTTTGAGTTGATAGTTATTAAAAAGAAAACAAAACA
AAAAAAAAAGGCAAGGCACAACAAAAAATATCTGGGCAATAAAAAAATATTTTAAACCAAAAAAAAAAAAAAAAAA

25

30

35

219

MEPSSKKLTGRLMLAVGGAVLGSQFGYNTGVINAPQKVEEFYNQTVWHRYGESILPTTLTLWLSLSVAIFSVGGMIGSFSVG
LFVNRFRNRNSMLMNLALFVSAVLMGFSKLGKSFEMLIIGRFLIGVYCLTTGFPVMYVGEVSPTAFRGALGTLHQLGIVVGI
LIAQVFGLDISMGNDLWPLLLSIIFIPALLQCIVLPFCPESPRFLLINRNEENRAKSVLKKLRGTADVTHDLQEMKEESRQMM
REKKVTILELFRSPAYRQPIILIAVVLQLSQQLSGINAVFYYSTSIPEKAGVQPPVYATIGSGIVNTAFTVVSFLFVVERAGRRTL
HLIGLAGMAGCAILMTIALALLEQLPWMSYLSIVAIFGFVAFVEVGPPIPFVFAELFSQGRPAIAVAGFSNWTNSNFIVGM
CFQYVEQLCGPYVFIIFTVLVLVFFIFTYFKVPETKGRFTDEIASGFRQGGASQSDKTPEELFHLPLGADSQV

45 220

TAGTCGCGGGTCCCCGAGTGAGCACGCCAGGGAGCAGGAGACCAACGACGGGGGTGCGAGTCAGAGTCGAGTGGGAGTCCCC
GGACCGGAGCACGAGCCTGAGCGGGAGAGCGCCGCTCGCACGCCGTCGCCACCCGCTACCCGGCGCAGCCAGAGCCACCAGC
GCAGCGTGCCATGGAGCCCAGCAGCAAGAAGCTGACGGGTGCGCTCATGCTGGCTGTGGGAGGAGCAGTGCCTGGCTCCCTGC
AGTTTGGGTACAACACTGGAGTCATCAATGCCCCCAGAAGGTGATCGAGGAGTTCTACAACCAGACATGGGTCCACCGCTATG
5 GGGAGAGCATCTGCCCCACCACGCTCACCACGCTCTGGTCCCTCTCAGTGGCCATCTTTCTGTGTGGGGCATGATTGGCTCCT
TCTCTGTGGGCTTTTCGTAAACCGCTTTGGCCGGCGGAATTCAATGCTGATGATGAACCTGCTGGCCCTTCGTGTCGCCGTGC
TCATGGGCTTCTCGAACTGGGCAAGTCTTTGAGATGCTGATCCTGGGCCGCTTCATCATCGGTGTGTACTGCCGCCGTGACCA
CAGGCTTCGTGCCCATGTATGTGGGTGAAGTGTCAACCCAGCCCTTCGTGGGGCCCTGGGCACCCTGCACCAGCTGGGCATCG
TCGTGCGCATCTCATCGCCAGGTGTTCCGCCCTGGACTCCATCATGGGCAACAAGGACCTGTGGCCCCGTGCTGTGAGCATCA
10 TCTTCATCCCGCCCTGCTGTCAGTGCATCGTGTGCCCTTCTGCCCCGAGAGTCCCGCTTCTGCTCATCAACCGCAACGAGG
AGAACCAGGGCAAGAGTGTGCTAAAGAAGCTGCGCGGGACAGCTGACGTGACCCATGACCTGCAGGAGATGAAGGAAGAGAGTC
GGCAGATGATGGGGAGAAGAAGTGCACCATCTGGAGTGTTCGCTCCCGCTCCCGCCCTACCGCCAGCCCATCCTCATCGTGTGG
TGCTGCAGCTGTCCAGCAGCTGTCTGGCATCAACGCTGTCTTCTATTACTCCACGAGCATCTTCGAGAAGCGGGGGTGCAGC
AGCTGTGTATGCCACCATTTGGCTCCGGTATCGTCAACACGGCCCTTCACTGTCTGTGCTGTGTGTGGTGGAGCGGAGCAGGCC
15 GCGGACCTGCACCTCATAGGCCCTCGCTGGCATGGCGGGTGTGCCATACTCATGACCATCGCGCTAGCACTGCTGGAGCAGC
TACCCTGGATGTCTATCTGAGCATCGTGGCCATCTTTGGCTTTGTGGCCTTCTTTGAAGTGGGTCTGGCCCCATCCCATGGT
TCATCGTGGCTGAACCTTCTCAGCCAGGGTCCACGTCCAGCTGCCATTGCCGTTCAGGGCTTCTCCAACGGACCTCAAATTTCA
TTGTGGGCATGTGCTTCCAGTATGTGGAGCAACTGTGTGGTCCCTACGTCTTCATCATCTTCACTGTGCTCCTGGTCTGTCT
TCATCTTCACTTACTTCAAAGTTCCTGAGACTAAAGGCCGACCTTCGATGAGATCGCTTCCGGCTTCCGGCAGGGGGAGCCA
20 GCCAAAGTGATAAGACACCCGAGGAGCTGTTCCATCCCTGGGGGTGATTTCCAAAGTGTGAGTCGCCCCAGATCACCAGCCG
GCCTGTCTCCAGCAGCCCTAAGGATCTCTCAGGAGCACAGGAGCTGGATGAGACTTCCAAACCTGACAGATGTACGCCGAGCC
GGGCTTGGGGCTCTTTCTCCAGCCAGCAATGATGTCCAGAAGAATATTTCAGGACTTAACGGCTCCAGGATTTTAAACAAAGCA
AGACTGTGTCTCAAATCTATTTCAGACAAGCAACAGGTTTATAAATTTTTTATTACTGATTTTGTATTATTATATCAGCCTGA
GTCTCCTGTGCCACATCCAGGCTTCACCTGAATGGTTCCATGCCTGAGGGTGGAGACTAAGCCCTGTGAGACACTTGCCT
25 TCTTCAACCAGCTAATCTGTAGGGCTGGACCTATGTCTTAAGGACACACTAATCGAATATGAACACAAAGCTTCTATCCAG
GAGGTGGCTATGGCCACCCGTTCTGTCTGGCTGGATCTCCCACTCTAGGGGTGAGGCTCCATTAGGATTTGCCCTTCCCATC
TCTTCTACCAACCACTCAAATTAATCTTTCTTTACCTGAGACCAGTTGGGAGCACTGGAGTGCAGGGAGGAGAGGGGAAGGG
CCAGTCTGGGCTGCCGGGTTCTAGTCTCCTTTGCACTGAGGGCCACACTATTACCATGAGAAGAGGGCTGTGGGAGCCTGCAA
ACTCAGTGTCTAAGAAGACATGGAGACTCCTGCCCTGTGTGTATAGATGCAAGATATTATATATATTTTGGTTGTCAATAT
30 TAAATACAGACATAAGTTATAGTATATCTGGACAAGCCAATCTGTAAATACACCACCTCACTCCTGTACTTACCTAAACAGA
TATAAATGGCTGGTTTTTAGAAACATGGTTTTGAATGCTTGTGGATTGAGGGTAGGAGGTTTGGATGGGAGTGAGACAGAAGT
AAGTGGGGTTGCAACCACTGCAACGGCTTAGACTTCGACTCAGGATCCAGTCCCTTACACGTACCTCTCATCAGTGTCTCTTG
CTCAAAAATCTGTTTGTATCCCTGTTCACCCAGAGAATATATACATTCTTTATCTTGACATTCAAGGCATTCTATCACATATTG
ATAGTTGGTGTCAAAAAAACACTAGTTTGTGCCAGCCGTGATGCTCAGGCTTGAAATCGCATTTATTTGAATGTGAAGGGAA
35 221
MFIMGLDPIPEELYEMLSDHSIRSFDLQRLHGDPEEDGAEIDLNMTRSHSGGELES LARGRRSLGSLTIAEPAMIAECKT
RTEVFEISRRLIDRTNANFLVWPPCEVQRCSGCCNRRNVQCRPTQVQLRPVQVRKIEIVRKKPIPKKATVTLEDHLACKCETV
AAARPVTRSPGGSQEQRAKTPQTRVTIRTVRVRPPKGRKFKHHTDKTALKETLGA
222
40 GTGTGTGTGTGTGTGCGCGCATCTGAGAGAGAGAGAGAGAGAGAGACTGACTGAGCAGGAATGGTGAGATGTTTATCATG
GGCCTCGGGGACCCATTTCCCGAGGAGCTTTATGAGATGCTGAGTGACCACTCGATCCGCTCCTTTGATGATCTCCAACGCTG
CTGCACGGAGACCCCGGAGAGGAAGATGGGGCCGAGTTGGACCTGAACATGACCCGCTCCCACTCTGGAGGCGAGCTGGAGAGC
TTGGCTCGTGGAAGAAGGAGCCTGGGTTCCTTGACCATTCCTGAGCCGCGCATGATCGCCGAGTGCAAGACGCGCACCGAGGTG
TTCGAGATCTCCCGGCGCTCATAGACCGCACCAACGCAACTTCTTGGTGTGGCGCCCTGTGTGGAGGTGCAGCGCTGCTCTC
45 GGCTGCTGCAACAACCGCAACGTGCACTGCGCCGCCACCCAGGTGCAAGCTGCGACCTGTCCAGGTGAGAAAGATCGAGATTGTG
CGGAAGAAGCCAATCTTTAAGAAGGCCACGGTGACGCTGGAAGACCACTGGCATGCAAGTGTGAGACAGTGGCAGCTGCACGG
CCTGTGACCCGAAGCCCGGGGGTTCCAGGAGCAGCGAGCCAAAACGCCCCAACTCGGGTGACCATTCGGACGGTGCAGGTC
CGCCGGCCCCCAAGGGCAAGCACCGGAAATTCAGCACACGCATGACAAGACGGCACTGAAGGAGACCTTGGAGCCTAGGGG
CATCGGCAGGAGAGTGTGTGGCAGGGTTATTTAATATGGTATTTGCTGTATTGCCCCATGGGGTCTTTGGAGTGATAATATT

5 GTTTCCTCGTCCGCTGTCTCGATGCCTGATTCCGACGGCCAATGGTGTCTCCCCACCCCTCCACGTGTCCGTCCACCCCTTC
CATCAGCGGGTCTCTCCAGCGGCTCCGGTCTTGCCAGCAGCTCAAAGAAGAAAAAGAGGACTGAACTCCATCGCCATCT
TCTTCCCTTAACCCAAGAACTTGGGATAAGAGTGTGAGAGAGACTGATGGGGTCGCTCTTTGGGGGAAACGGGTTCCTTCCCC
TGACCTGGCCTGGGCCACACCTGAGCGCTGTGGACTGTCTGAGGAGCCCTGAGGACCTCTCAGCATAGCCTGCCTGATCCCT
10 GAACCCCTGGCCAGCTCTGAGGGGAGGCACCTCCAGGCAGGCCAGGCTGCCTCGGACTCCATGGCTAAGACCACAGACGGGCAC
ACAGACTGGAGAAAACCCCTCCACGGTGGCCAAACACCACTGACCTCGTCTCCCTGGTGCCTCTGTGCACAGTGGCTTCTTTT
CGTTTTCGTTTGAAGACGTGGACTCCTCTTGGTGGGTGTGGCCAGCACACCAAGTGGCTGGGTGCCCTCTCAGGTGGGTAGA
GATGGAGTTTGTCTGTGAGGTGGTGTAGATGGTGACCTGGGTATCCCTGCCTCCTGCCACCCCTTCTCCCCATACTCCACTC
TGATTACCTCTTCTCTGGTTCCTTTCATCTCTTACCTCCACCTGCATTTTCTCTTGTCTGGCCCTTCAGTCTGCTCCA
15 CCAAGGGGCTCTTGAACCCCTTATTAAGGCCCCAGATGACCCAGTCACTCTCTTAGGGCAGAAGACTAGAGGCCAGGGCAG
CAAGGGACCTGCTCATCATATTCCAACCCAGCCAGACTGCCATGTAAGGTGTGTCAGGGTGTGTACTGCACAAGGACATTGTA
TGCAGGGAGCACTGTTTACATCATAGATAAAGCTGATTTGTATATTTATGACAATTTCTGGCAGATGATAGGTAAAGAGGAA
AAGGATCCTTTTCTAATTCACACAAAGACTCCTTGTGGACTGGCTGTGCCCTGATGCAGCCTGTGGCTGGAGGCCAAATA
GGAGGGAGACTGTGGTAGGGCAGGGAGGCAACTGCTGTCCACATGACCTCCATTTCCCAAAGTCTCTGCTCCAGCAACTG
20 CCCTTCCAGGTGGGTGTGGGACACCTGGGAGAAGGTCTCCAAGGAGGGTGCAGCCCTCTTGCCCGCACCCCTCCCTGCTTGCA
CACTTCCCCATCTTTGATCCTTCTGAGCTCCACCTCTGGTGGCTCCTCTAGGAAACAGCTCGTGGGTGGGAATGGGGGAGA
GAAGGGAAAAGATCCCCAAGACCCCTGGGGTGGGATCTGAGCTCCACCTCCCTTCCACCTACTGCATTTCCCCCTTCCCG
CCTTCCAAAACCTGCTTCCCTCAGTTTGTAAAGTCGGTATTATATTTTGGGGCTTCTCTTTTATTTTAAATGTAAATTT
TATTTATATTTCCGTATTTAAAGTTGT

20 223

MATLKDQLIYNLLKEEQTPQNKITVVGAVGMACAISILMKDLADELALVDVIEDKLKEMMDLQHSLSFLRTPKIVSGKDYD
VTANSKLVIIITAGARQQEGESRLNLVQRNVNIFKPIIPNVVKYSPNCKLLIVSNPVDILTYVAVWKISGFPKNRVIGSGCNLDSA
RFRYLMGERLGVHPLSCHGWVLGEHGDSSVPVWSGMNVAGVSLKTLHPDLGTDKKEQWKEVHKQVVEAYEVIKLKGYTSWAI
GLSVADLAESIMKNLRRVHPVSTMIKGLYGIKDDVFLSVPCILQNGISDLVKVTLTSEEARLKSADTLWGIQKELQF

25 224

TGCTGCAGCCGCTGCCCGGATTCCGGATCTCATTGCCACGCGCCCCGACGACCGCCCGACGTGCATTCCCGATTCTCTTTTGG
TTCCAAGTCCAATATGGCAACTCTAAAGGATCAGCTGATTATATAATCTTCTAAAGGAAGAACAGACCCCCAGAATAAGATTAC
AGTTGTGGGGTGGTGTCTTGGCATGGCCTGTGCCATCAGTATCTTAATGAAGGACTTGGCAGATGAACTTGCTCTTGTGA
TGTCATCGAAGACAAATTGAAGGGAGAGATGATGGATCTCCAACATGGCAGCCTTTTCTTAGAACACCAAAGATTGTCTCTGG
30 CAAAGACTATAATGTAACGCAAACTCCAAGCTGGTCAATATCAGCGCTGGGGCAGCTCAGCAAGAGGGAGAAAGCCGTCTTAA
TTTGGTCCAGCGTAACGTGAACATATTTAAATTCATCATTTCTTAATGTTGTAAATAACAGCCCGAACTGCAAGTTGCTTATTGT
TTCAATCCAGTGGATATCTTGACCTACGTGGCTTGGGAAGATAAGTGGTTTCCCAAAAACCGTGTATTGGAAGTGGTTGCAA
TCTGGATTACGCCGATTCCGTACCTGATGGGGAAAGGCTGGGAGTTACCCATTAAAGCTGTCTGATGGGTGGGTCTTGGGGA
ACATGGAGATTCCAGTGTGCTGTATGGAGTGGAAATGAATGTTGCTGGTGTCTCTGAAAGACTCTGCACCCAGATTTAGGGAC
35 TGATAAAGATAAGGAACAGTGGAAAGAGGTTTCAAGCAGGTGGTGTAGAGTGTCTATGAGGTGATCAAACTCAAAGGCTACAC
ATCCTGGGCTATTGGACTCTCTGTAGCAGATTTGGCAGAGAGTATAATGAAGATCTTAGGCGGGTGACCCAGTTTCCACCAT
GATTAAGGGTCTTTACGGAATAAAGGATGATGTCTTCTTAGTGTCTCTTGCAATTTGGGACAGAATGGAATCTCAGACCTTGT
GAAGGTGACTCTGACTTCTGAGGAAGAGGCCGTTTGAAGAAGACTGCAGATACACTTTGGGGGATCCAAAAGGAGCTGCAATT
TTAAAGTCTTCTGATGTATATCATTTTCACTGTCTAGGCTACAACAGGATTTAGGTGGAGGTTGTGATGTTGCTCTTTTAT
40 CTGATCTGTGATTAAAGCAGTAATATTTTAAAGATGGACTGGGAAAAACATCAACTCCTGAAGTTAGAAATAAGAAATGGTTGTGA
AAATCCACAGCTATATCTGTATGCTGGATGGTATTAATCTTGTGTAGTCTTCAACTGGTTAGTGTGAAATAGTTCTGCCACCTC
TGACGCCACCACTGCCAATGCTGTACGTACTGCATTTGCCCTTGGAGCCAGGTGGATGTTTACCGTGTGTATATAACTTCTCTGG
CTCCTTCACTGAACATGCCCTAGTCCAACATTTTTCAGTGAAGTATATCAGTGTATACATTACCATATAATGTAAAGATCTACATA
TTGTGCATAATTTCTTCAAAGGATCTTATTTTGAAGTATATCAGTGTATACATTACCATATAATGTAAAGATCTACATA
45 CAAACAATGCAACCAACTATCCAAGTGTATACCAACTAAAACCCCAATAAACCTTGAACAGT

225

METPAWPRVPRPETAVARTLLLGWVFAQVAGASGTTNTVAAYNLTKSTNFKTILEWEPKPVNQVYTVQISTKSGDWKSKCFYT
TDTECDLTDEIVKDVKTLYLARVFSYPAGNVSTGSAGEPLYENSPEFTPYLETNLGQPTIQSFEQVGTKNVTVEDERTLVRR

NNTFLSLRDVFGKDLIYTLYYWKSSSSGKKTAKTNTNEFLIDVDKGENYCFVQAVIPSRVTNKRKSTDSPVECMGQEKGEFREI
FYIIGAVVVFVVIILVILAILSLHKCRKAGVGQSWKENSPLNV5

226

5 AAGACTGCGAGCTCCCCGCACCCCTCGCACTCCCTCTGGCCGGCCAGGGCGCCTTCAGCCCCAACCCTCCCCAGCCCCACGGGC
GCCACGGAACCCGCTCGATCTCGCCGCCAACTGGTAGACATGGAGACCCCTGCCCTGGCCCCGGGTCCCGCGCCCCGAGACCGCC
GTCGCTCGGACGCTCTGCTCGGCTGGGTCTTCGCCCAGGTGGCCGGCGCTTCAGGCACACAAAATACTGTGGCAGCATATAAT
TTAACTTGGAAATCAACTAATTTCAAGACAATTTTGGAGTGGGAACCCAAACCCGTCATCAAGTCTACACTGTTCAAATAAGC
ACTAAGTCAGGAGATTGGAAAAGCAAATGCTTTTACACAACAGACACAGAGTGTGACCTCACCGACGAGATTGTGAAGGATGTG
AAGCAGACGTACTTGGCAGGGGTCTTCTCCTACCCGGCAGGGAATGTGGAGAGCACCGGTTCTGCTGGGAGCCCTCTGTATGAG
10 AACTCCCCAGAGTTACACCTTACCTGGAGACAAACCTCGGACAGCCAAACAATTCAGAGTTTGAACAGGTGGGAACAAAAGTG
AATGTGACCGTAGAAGATGAACGGACTTTAGTCAGAAGGAACAACACTTTCCTAAGCCTCCGGGATGTTTTTGGCAAGGACTTA
ATTTATACACTTTATTATTGGAATCTTCAAGTTTCAGGAAGAGAAAACAGCCAAAACAACTAATGAGTTTTTGGATTGATGTG
GATAAAGGAGAAAATACTGTTTTAGTGTTCAGCAGTGATTCCCTCCCGAACAGTTAACCGGAAGAGTACAGACAGCCCGGTA
GAGTGTATGGGCCAGGAGAAAAGGGAATTCAGAGAAATATTCTACATCATTTGGAGCTGTGGTATTTGTGGTCATCATCTTGTCT
15 ATCATCTGGCTATATCTCTACACAAGTGTAGAAAGGAGGAGTGGGGCAGAGCTGGAAGGAGAACTCCCCACTGAATGTTTCA
TAAAGGAAGCACTGTTGGAGCTACTGCAATGCTATATTGCACGTGTGACCGAGAACTTTTAAGAGGATAGAATACATGGAAACG
CAATGAGTATTTTCGGAGCATGAAGACCCCTGGAGTTCAAAAACTCTTGATATGACCTGTTATTACCATTAGCATTCTGGTTTT
GACATCAGCATTAGTCACTTTGAATGTAACGAATGGTACTACAACCAATTCCAAGTTTTAATTTTTAACACCATGGCACCTTT
TGCACATAACATGCTTTAGATTATATATTCGCCACTTAAGGATTAACCGGTCGTCACGCAAAAAACAATGGGAAAATGTCTT
20 AAAAAATCTGGGTGGACTTTTGAAGCTTTTTTTTTTTTTTTTTTTTGGAGACGGAGTCTTGCTCTGTGCCCAGGCTGGAGT
GCAGTAGCACGATCTCGGCTCACTTGCACCCCTCGTCTCTCGGGTTCAAGCAATGTCTGCCTCAGCCTCCCGAGTAGCTGGGA
TTACAGGTGCGCACTACCACGCCAAGCTAATTTTTGTATTTTTTAGTAGAGATGGGGTTTACCATCTTGGCCAGGCTGGTCTT
GAATTCCTGACCTCAGTGATCCACCCACCTTGGCCTCCCAAAGATGCTAGTATTATGGGCGTGAACCACCATGCCCAGCCGAAA
AGCTTTTGGGGGTGACTTCAATCCATGTAGGAAAGTAAATGGAAGGAAATTTGGGTGCATTTCTAGGACTTTTCTAACATAT
25 GTCTATAATATAGTGTTTAGGTTCTTTTTTTTTTTCAGGAATACATTTGGAAATTCAAAACAATTTGGGCAAACTTTGTATTAAATG
TGTTAAGTGCAGGAGACATTGGTATTCTGGGCAGCTTCCTAATATGCTTTACAATCTGCACTTTAACTGACTTAAGTGGCATT
AACATTTGAGAGCTAACTATATTTTTATAAGACTACTATACAACTACAGAGTTATGATTTAAGGTACTTAAAGCTTCTATGG
TTGACATTGTATATATAATTTTTTAAAGGTTTTTCTATATGGGATTTTCTATTTATGTAGGTAATATTGTCTATTTGTAT
ATATTGAGATAAATTTATTTAATATACTTTAAATAAAGGTGACTGGGAATTGTT

30 227

MNFLLSWVHWSLALLLYLHHAKWSQAAPMAEGGQNHHEVVKFMVYQRSYCHPIETLVDIFQEYPDEIEYIFKPSCVPLMRG
GCCNDEGLECVPTESNITMQIMRIKPHQGHIGEMSFLOHNKCECRPKKDRARQENPCGPCSERRKHLFVQDPQTKCSCKNT
DSRCKARQLELNERTCRCDKPRR

228

35 TCGGGCCTCCGAAACCATGAACTTTCTGCTGTCTTGGGTGCATTGGAGCCTTGCCCTGCTGCTCTACCTCCACCATGCCAAGTG
GTCCCAGGTGCACCCATGGCAGAAAGGAGGGGAGGAGATCATCACGAAGTGGTGAAGTTTCATGGATGTCTATCAGCGCAGCTA
CTGCCATCCAATCGAGACCCCTGGTGGACATCTTCCAGGAGTACCCTGATGAGATCGAGTACATCTTCAAGCCATCCTGTGTGCC
CCTGATGCGATGCGGGGGCTGTGCAATGACGAGGGCCTGGAGTGTGTGCCACTGAGGAGTCCAACATCACCATGCAGATTAT
GCGGATCAAACCTCACCAAGGCCAGCACATAGGAGAGATGAGCTTCTTACAGCAACAACAAATGTGAATGCAGACCAAGAAAGA
40 TAGAGCAAGACAAGAAAATCCCTGTGGGCCTTGCTCAGAGCGGAGAAAGCATTTGTTTGTACAAGATCCGCAGACGTGTAAATG
TTCCTGCAAAAACAGACTCGCGTTGCAAGGCGAGGAGCTTGAGTTAAACGAACGTACTTGCAGATGTGACAAGCCGAGGCG
GTGAGCCGGCAGGAGGAAGGAGCCTCCCTCAGCGTTTCGGGAACAGATCTCTCACCAGG

229

MSREMQDVLAEVKPLVEKGETITGLLQEFVQEQDIETLHGSVHVTLCGTPKGNRPVILTYHDIGMNHKTCYNPLFNYEDMQE
45 ITQHFVAVCHVDAPGQDGAASFPAGYMPSPMDQLAEMLPVQLQFGLKSIIGMTGAGAYTLTRFALNNPEMVEGLVLINVNPC
AEGWMDWAASKISGWTQALPDMVVSHLFGKEEMQSNVEVVHTYRQHIVNDMNPGLHLFINAYNSRRDLEIERPMPGHTHTVTLQ
CPALLVVGDSPPAVDAVVECNKLDPTKTLLKMACDGLPQISQPAKLAEAFKYFVQGMGYMPSASMTRLMRSRTASGSSVTS
LDGTRSRSHSTEGTRSRSHSTEGTRSRSHSTEGAHLDITPNSGAAGNSAGPKSMEVSC

230

TGAAGCTCGTCAGTTCACCATCCGCCCTCGGCTTCCGCGGGCGCTGGGCCGCCAGCCTCGGCACCGTCCTTTCTTTCTCCCT
CGCGTTAGGCAGGTGACAGCAGGGACATGTCTCGGGAGATGCAGGATGTAGACCTCGCTGAGGTGAAGCCTTTGGTGGAGAAAG
GGGAGACCATCACCGCCCTCTGCAAGAGTTTGATGTCCAGGAGCAGGACATCGAGACTTTACATGGCTCTGTTACGTCACGC
5 TGTGTGGGACTCCCAAGGGAACCGGCCGTGTCATCCTACCTACCATGACATCGGCATGAACCACAAAACCTGCTACAACCCCC
TCTTCAACTACGAGGACATGCAGGAGATCACCCAGCACTTTGCCGTCTGCCACGTGGACGCCCTGGCCAGCAGGACGGCGCAG
CCTCCTTCCCCCGAGGGTACATGTACCCCTCCATGGATCAGCTGGCTGAAATGCTTCTGGAGTCTTCAACAGTTTGGGCTGA
AAAGCATATTGGCATGGGAACAGGAGCAGGCGCTACACCTAACTCGATTGCTCTAAACAACCTGAGATGGTGGAGGGCC
TTGTCTTATCAACGTGAACCTTGTGCGGAAGGCTGGATGGACTGGGCCGCCCTCAAGATCTCAGGATGGACCAAGCTCTGC
10 CGGACATGGTGGTGTCCACCTTTTGGGAAGGAAGAAATGCAGAGTAACGTGGAAGTGGTCCACACCTACGCCAGCACATTG
TGAATGACATGAACCCCGCAACCTGCACCTGTTTCATCAATGCCACACAGCCGGCGCGACCTGGAGATTGAGCGACCAATGC
CGGGAACCCACACAGTACCCCTGCAGTGGCTGCTCTGTTGGTGGTGGGGACAGCTCGCTGCAGTGGATGCCGTGGTGGAGT
GCAACTCAAAATTGGACCAACAAGACCACTCTCTCAAGATGGCGGACTGTGGCGGCTCCCGCAGATCTCCAGCCGGCCA
AGCTCGCTGAGGCCCTTCAAGTACTTCGTGCAGGGCATGGGATACATGCCCTCGGCTAGCATGACCCGCTGATGCGGTCCCGCA
15 CAGCTCTGTTCCAGCGTCACTTCTCTGGATGGCACCAGCCGCTCCACACAGCAGGAGGCCACCCGAAGCCGCTCCCA
CCAGCGAGGGCACCAGCGAGCCGCTCGCACACAGCAGGAGGGGCCACCTGGACATCACCCCAACTCGGGTGTCTGGGAACA
GCGCGGGGCCAAGTCCATGGAGGTCTCTGCTAGGCGGCTGCCAGCTGCCGCCCGGACTCTGATCTCTGTAGTGGCCCC
CTCTCCCCCGGCCCTTTTCGCCCCCTGCCTGCCATACTGCGCCTAACTCGGTATTAATCCAAAGCTTATTTTGTAAAGTGTAG
CTCTGGTGGAGACAAATGAGGTCTATTACGTGGGTGCCCTCTCCAAAGGCGGGTGGCGGTGGACCAAGGAAGGAAGCAAGCA
20 TCTCCGCATCGCATCTCTTCCATTAACAGTGGCGGTTGCCACTCTCTCCCTCCCTCAGAGACACCAAACTGCCAAAAAC
AAGACGCGTAGCAGCACACACTTCACAAAGCCAAGCCTAGGCCGCCCTGAGCATCTGGTTCAAACGGGTGCCGTGAGAAGG
CCAGCCGCCCACTTCCCGTTTCTCTTTAACTGAGGAGAAGCTGATCCAGCTTTCCGGAACAAAATCTTTTCTCTCATTTGGG
GAGGGGGTAATAGTGACATGCAGGCACCTCTTTTAAACAGGCAAAACAGGAAGGGGAAAAGGTGGGATTGATGTCGAGGCTA
GAGGCATTTGGAACAACAAATCTACGTAGTTAACTTGAAGAAACCGATTTTTAAAGTTGGTGCATCTAGAAAGCTTTGAATGCA
25 GAAGCAAAACAGCTTGAATTTTCTAGCATCTCTTAATGTGCAGCAAAAGCAGGCAACAAAATCTCTGGCTTTACAGACAAAA
ATATTTACAGCAACGTTGGGCATCATGGTTTTTGAAGGCTTTAGTTCTGCTTTCTGCTCTCTCCACAGCCCCAACCTCCAC
CCCTGATACATGAGCCAGTGATATTCTTGTTCAGGGAGAAGATCATTTAGATTTGTTTTGCATTCCTTAGAATGGAGGGCAAC
ATTCACAGCTGCCCTGGCTGTGATGAGTGTCTTGCAGGGGCCGAGTAGGAGCACTGGGGTGGGGGCGGAATTTGGGGTACT
CGATGTAAGGGATTCTTTGTTGTTGTGTGATCCAGTGCAGTTGTGATTTCTGTGGATCCAGCTTGGTCCAGGAATTTTGA
30 GAGATTGGCTTAAATCCAGTTTTCATCTTCGACAGCTGGGCTGGAACGTGAATCAGTAGCTGAACCTGTCTGACCCGGTCAC
GTTCTTGGATCCTCAGAACTCTTTGCTCTTGTGCGGGTGGGGTGGGAACCTACGTGGGGAGCGGTGGCTGAGAAAATGTAAGG
ATTCGTGGAATACATATTCATGGACTTTCTCTCTCTGCTCTCTGCTTCTCTCTTTCTGCTCTCTTAACTTTGCGCGAATGGGCA
GACAAACACTGACGTTTCTGGGTGGCCAGTGGCGCTGCCAGGTTCTCTGTACTACTGCTTGTACTTTTCATTTTGGCTCACCGT
GGATTTTCTCATAGGAAGTTTGGTCAGAGTGAATGAATATTGTAAGTCAGCCACTGGGACCCGAGGATTTCTGGGACCCGCA
35 GTTGGGAGGAGGAAGTAGTCCAGCCTTCCAGGTGGGCGTGAGAGGAATGACTCGTTACCTGCCGCCCATCACCTTGGAGGCCCT
TCCCTGGCCTTGTAGTAGAAAAGTCGGGGATCGGGGCAAGAGAGGCTGAGTACGGATGGGAACTATTGTGCACAAGTCTTTCCA
GAGGAGTTTCTTAATGAGATATTGTATTTATTCCAGACCAATAAATTTGTAACCTTGCAAAAAAAAAAAAAAAAAAAAA

231

MIWYILIIGILLPQSLAHPGFFTSIGQMTDLIHTKDLVTSKDYIKAEBDKLEQIKKWAEEKLDRLTSTATKDPEGFVGHVPNA
40 FKLKRLNTEWSELENLVLKDMDSGFISNLTIQRPVLSNDEDQVGAALKLLRLQDTYNLDTDTISKGNLPGVKHKSFLTAEEDCF
ELGKVAYTEADYYHTELWMEQALRQLDEGEISTIDKVSVDLYLSYAVYQQGDLDKALLLTKKLELDPEHQORANGLKYFEYIM
AKEKDVNKSASDDQSDQKTTPKKGVAVDYLPERQKYEMLCRGEIGKMTPRRQKKLFCRYHDGNRNPKFILAPAKQEDENDKPR
IIRFHDIISDAEIEIVKDLAKPRLSRATVHDPETGKLTTAQYRVSKSAWLSGYENPVVSRINMRIQDLTGLDVTAEELQVANY
GVGGQYEPHFDPARKDEPDAFKELGTGNRIATWLFYMSDVSAGGATVFPVGVASVWPKKGTAVFWYNLFASGEGDYSTRHAACP
45 VLVGKWNKSNKWLHERGQEFRRPCTLSELE

232

GAGCGGGCTGAGGGTAGGAAGTAGCCGCTCCGAGTGGAGGCGACTGGGGCTGAAGAGCGCGCGCCCTCTCGTCCCACCTTTCC
AGGTGTGTGATCTGTAAAAATTAAATCTTCCAAGATGATCTGGTATATATTAATTATAGGAATTCTGCTTCCCCAGTCTTTGGC
TCATCCAGGCTTTTTFACCTCAATTGGTCAGATGACTGATTTGATCCATACTGAGAAAGATCTGGTGAATTTCTCTGAAAGATTA
50 TATTAAGGCAGAAGAGGACAAGTTAGAACAATAAAAAATGGGCAGAGAAGTTAGATCGGCTAACTAGTACAGCGACAAAAGA

- TCCAGAAGGATTTGTTGGGCATCCAGTAAATGCATTCAAATTAATGAAACGCTCTGAATACTGAGTGGAGTGAGTTGGAGAATCT
GGTCCTTAAGGATATGTCAGATGGCTTTATCTCTAACCTAACCATTCAGAGACCAGTACTTTCTAATGATGAAGATCAGGTTGG
GGCAGCCAAAGCTCTGTTACGCTCTCCAGGATACCTACAATTTGGATACAGATACCATCTCAAAGGGTAATCTTCCAGGAGTGAA
ACACAAATCTTTCTAACGGCTGAGGACTGCTTTGAGTTGGGCAAAGTGGCCTATACAGAAGCAGATTATTACCATACGGAACCT
5 GTGGATGGAACAAGCCCTAAGGCAACTGGATGAAGGCGAGATTTCTACCATAGATAAAGTCTCTGTTCTAGATTATTTGAGCTA
TGCGGTATATCAGCAGGGAGACCTGGATAAGGCACTTTTGCTCACAAGAAGCTTCTTGAACATGATCTGAACATCAGAGAGC
TAATGGTAACTTAAATATTTGAGTATATAATGGCTAAAGAAAAAGATGTCAATAAGTCTGCTTCAGATGACCAATCTGATCA
GAAAACCTACACCAAGAAAAAGGGGTTGCTGTGGATTACCTGCCAGAGAGACAGAAGTACGAAATGCTGTGCCGTGGGGAGGG
TATCAAAATGACCCCTCGGAGACAGAAAAACTCTTTTGCCGCTACCATGATGGAAACCGTAATCCTAAATTTATTTCTGGCTCC
10 AGCTAAACAGGAGGATGAATGGGACAAGCCTCGTATTATTCGCTTCCATGATATTTATTTCTGATGCAGAAATGAAATCGTCAA
AGACCTAGCAAAACCAAGGCTGAGCCGAGCTACAGTACATGACCCTGAGACTGGAAAATTGACCACAGCACAGTACAGAGTATC
TAAGAGTGCCTGGCTCTCTGGCTATGAAAATCCTGTGGTGTCTCGAATTAATATGAGAATACAAGATCTAACAGGACTAGATGT
TTCCACAGCAGAGGAATTACAGGTAGCAAATTAAGAGTTGGAGGACAGTATGAACCCCATTTTGACTTTGCACGGAAGATGA
GCCAGATGCTTTTCAAAGAGCTGGGACAGGAAATAGAATTGCTACATGGCTGTTTATATGAGTGTGTCTGCAGGAGGAGC
15 CACTGTTTTTCTGAAGTTGGAGCTAGTGTTTGGCCCAAAAAAGGAACCTGCTGTTTTCTGGTATAATCTGTTTGGCAGTGGAGA
AGGAGATTATAGTACACGGCATGCAGCCTGTCCAGTGTAGTTGGCAACAAATGGGTATCCAATAAATGGCTCCATGAACGTGG
ACAAGAATTTGGAAGACCTTGTACGTTGTCAGAATTGGAATGACAAACAGGCTTCCCTTTTCTCCTATTGTTGTAATCTTATG
TGCTGATATACACATTTCCATAGTCTTAACCTTTCAGGAGTTTACAATTGACTAACACTCCATGATTGATTCAGTCATGAACCT
CATCCCATGTTTTCATCTGTGGACAATTGCTTACTTTGTGGGTTCTTTTAAAGTAACACGAAATCATCATATTGCATAAAACCT
20 TAAAGTTCTGTTGGTATCACAGAAGACAAGGCAGAGTTTAAAGTGAGGAATTTTATATTTAAAGAACTTTTGGTTGGATAAAA
ACATAATTTGAGCATCCAGTTTATGATTTTCACTACATCTCAGTTGGTGGGTGTTAAGCTAGAATGGGCTGTGTGATAGGAAAC
AAATGCCCTACAGATGTGCTAGGTGTTCTGTTTACCTAGTGTCTTACTCTGTTTCTGGATCTGAAGACTAGTAATAAAGTAG
GACACTAACTGGGTTCCATGTGATTGCCCTTTCATATGATCTTCTAAGTTGATTTTTTCTCCCAAGCTCTTTTAAAGAAAG
TATACTGTATTTTACCAACCCCTCTCTTTTCTTTTAGCTCCTCTGTGGTGAATTAACGCTACTTGAGTTAAAAATATTTGATT
25 TTTTTTTTTTTTTAATGAAAGTCTGCTGATAACAACTGGGCCCTTCTTAACATAAATGCTCACCACCTAGCCCTGTTTTTTTA
TCCCTTTTTTAAAAATGACAGATGATTGTTTCAGGAATTTGCTGTTTTTCTTAGTGCTAATACCTTGCCCTCTTATTCCTGCTA
CAGCAGGGTGGTAATATTGGCATTTCTGATTAAATACTGTGCCTTAGGAGACTGGAAGTTAAAAATGTACAAGTCTTTTCAGTG
ATGAGGGAATTGATTTTTTTTTAAAGTCTTTTTCTTAGAAAGCCAAATGTTTGTTTTTTAAAGATTCTGAAATGTGTTGTGAC
AACAAATGACCTATTATGATCTTAAATCTTTTTT
- 30 233
- MSQNGAPGMQEEESLQGSWVELHFSNNGNGSVPASVSIYNGDMEXILLDAQHESGRSSSKSSHCDSPPRSQTPOQDINRASETDT
HSIGEKNSQSEBDDIERRKEVESILKKNSDWIWDSSRPENIPKEFLFKHPKRTATLSMRNTSVMKKGGIFSAEFLKVFLPS
LLLSHLLAIGLGIYIGRRLTTSTSTF
- 234
- 35 CCTCCGCTCAGTCCGGGAGCGCACGTGGGCCGCGCGCTCCGACCTCCGCTTTCCACCGCCCGCAGCTGAAGCACATCCCGCA
GCCCCGGCGCGGACTCCGATCGCCGAGTTGCCCTCTGGCGCCATGTCGAGAACGGAGCGCCCGGGATGCAGGAGGAGAGCCTG
CAGGGCTCCTGGGTAGAACTGCACCTCAGCAATAATGGGAACGGGGGACGCGTTCCAGCCTCGGTTTCTATTTAATGGAGAC
ATGGAAAAAATACTGCTGGACGCACAGCATGAGTCTGGACGGAGTAGCTCCAAGAGCTCTCACTGTGACAGCCACCTCGCTCG
CAGACACCACAAGATACCAACAGGGCTTCTGAAACAGATACCCATAGCATTTGGAGAGAAAAACAGCTCAGAGTCTGAGGAAGAT
40 GATATTGAAAGAAGGAAAGAGTTGAAAGCATCTTGAAGAAAACTCAGATTGGATATGGGATTGGTCAAGTCGGCCGGAAAAAT
ATTCCCCCAAGGAGTTCTCTTTAAACACCCGAAGCGCACGGCCACCCCTCAGCATGAGGAACACGAGCGTCATGAAGAAAGGG
GGCATATTTCTGCAAGATTTCTGAAAGTTTCTTCCATCTCTGCTGCTCTCTCATTTGCTGGCCATCGGATTGGGGATCTAT
ATTGGAAGGCGTCTGACAACCTCCACCAGCACCTTTTGATGAAGAACTGGAGTCTGACTTGGTTTCGTTAGTGGATTACTTCTGA
GCTTGCAACATAGCTCACTGAAGAGCTGTAGATCTGCGGTGGCCACGTCACCTTGTGTTTATTTGTTCTGTAAATGCTGCGTT
45 CCTAATTTAGTAAAAATAAAGAATAGACACTAAATCATGTTGATCTATAATTACACCTATGGGATCAATAAGCATGTCAGACT
GATTAATGTCTACTGTGAAAAATTTGCTAGTAAATTTCAATTTGATATTAGATATAAATATCTGAATATAAATAAATTTAATATA
CTAGTCATGATGTGTGTTGTTTAAAAATATCTGCAACCTTAATTCAGCTGAAGTACTTTATATTTCAAAGAATGAATAA
CATTGATAATAAATCGCTACTTTAAGGGGTTTGTCCAAAAATAAATATTGTGGCCTTATATATCACTATTGTAGAAAGTATT
ATTTAATTTAAATGGATGCAGGTTGTCTACTAAAGAAAGATTATATATAACTATGCTAATTTGTTTATAATCAACAGAAACCAAG
50 ATAGAGCTACAACTCAGCTGTACAGTTCTGTACACTAAACTCTTCTTGCTTTTGCAATTATAAGGAATTAAGTCTCCGATTATTA

GGTGATCACCCCTGGATGATCAGTTTTCTGCTGAAGGCACCTACTCAGTATCTTTTCCTCTTTATCACTCTGCATTGGTGAATTT
AATCCTCTCCTTTGTGTCAACTTTTGTGTGCTTTTAAATCAGCTTTATTCTAAGCAAATCTGTGTCTACTTTAAAAAAGTGG
AAATGGAAAAAAAATAAATCTT

5 235

MQMSPALTCLVLGLALVFEGESAVHPPSYVAHLASDFGVRVFQOVAQASKDRNVVFSYGVASVLAMLQLTTGGETQQQIQAA
MGFKIDDKGMAPALRHLKELMGPWNKDEISTDAIFVQRDLKLVLQGFMPHFRLFRSTVKQVDFSEVERARFIINDWVKTHTK
GMISNLLKGAVDQLTRLVLVNALYFNGQWKTPFPDSSTHRRLFHKSDGSTVSVPMMAQTNKFNYTEFTTPDGHYYDILELPYH
GDTLSMFIAAPYEKEVPLSALTNILSAQLISHWKGNMTRLPRLLVLPKFSLETEVDLRKPLENLGMTDMFRQFQADFTSLSDQE
10 PLHVAQALQKVKIEVNESGTVASSSTAVIVSARMAPEEIIIMDRPFLFVVRHNFTGTVLFMGQVMEP

236

GAATTCCTGCAGCTCAGCAGCCCGCCGAGAGCAGGACGAACCGCCAATCGCAAGGCACCTCTGAGAACTTCAGGATGCAGATG
TCTCCAGCCCTCACCTGCCTAGTCTGGGCTGGCCCTTGTCTTTGGTGAAGGGTCTGCTGTGCACCATCCCCCATCTTACGTG
GCCCACCTGGCCTCAGACTTCGGGGTGAAGGTGTTTTCAGCAGGTGGCGCAGGCCTCCAAGGACCGCAACGTGGTTTCTCACCC
15 TATGGGGTGGCCTCGGTGTTGGCCATGCTCCAGCTGACAACAGGAGGAGAAACCCAGCAGCAGATTCAAGCAGCTATGGGATTC
AAGATTGATGACAAGGGCATGGCCCCCGCCCTCCGGCATCTGTACAAGGAGCTCATGGGGCCATGGAACAAGGATGAGATCAGC
ACCACAGACGCGATCTTCGTCCAGCGGATCTGAAGCTGGTCCAGGGCTTCATGCCCACTTCTTCAGGCTGTTCCGGAGCAGC
GTCAAGCAAGTGGACTTTTCAGAGGTGGAGAGAGCCAGATTCATCATCAATGACTGGGTGAAGACACACACAAAAGGTATGATC
AGCAACTTGCTTGGGAAAGGAGCCGTGGACCAGCTGACACGGCTGGTCTGGTGAATGCCCTTACTTCAACGGCCAGTGGGAAG
20 ACTCCCTTCCCCGACTCCAGCACCCACCGCCGCTCTTCCACAAATCAGACGGCAGCACTGTCTCTGTGCCCCATGATGGCTCAG
ACCAACAAGTTCACTATACTGAGTTTCAACACGCGCCGATGGCCATTACTACGACATCTTGGAACTGCCTTACCACGGGGACACC
CTCAGCATGTTTCAATGCTGCCCCCTTATGAAAAAGAGGTGCTCTCTCTGCCCCACCAACATTCTGAGTGCCCCAGCTCATCAGC
CACTGGAAGGCAACATGACCAGGCTGCCCGGCTCTCGGTTCTGCCCAAGTTCTCCCTGGAGACTGAAGTCGACCTCAGGAAG
CCCCTAGAGAACCTGGGAATGACCGACATGTTTCAGACAGTTTCAGGCTGACTTCACGAGTCTTTTCAGACCAAGAGCCTCTCCAC
25 GTCGCGCAGGCGCTGCAGAAAGTGAAGATCGAGGTGAACGAGAGTGGCAGCGTGGCCTCCTCATCCACAGCTGTCATAGTCTCA
GCCCCATGCCCCCGAGGAGATCATCATGGACAGACCTTCTCTTTGTGGTCCGGCACAACCCACAGGAACAGTCTTTTC
ATGGGCCAAGTGAAGAACCTGACCTGGGGAAAGACGCTTCACTGGGACAAAACCTGGAGATGCATCGGAAAGAAAGAAAC
TCCGAAGAAAAGAAATTTAGTGTAAATGACTCTTTCTGAAGGAAGAGAAGACATTTGCCCTTTTGTAAAAGATGGTAAACCAGA
TCTGTCTCCAAGACCTTGGCCTCTCCTTGGAGGACCTTTAGGTCAAACCTCCCTAGTCTCCACCTGAGACCCCTGGGAGAGAAGTT
30 TGAAGCACAACTCCCTTAAGGTCTCCAAACCAGACGCTGACGCTGCGGGACCATCTGGGGACCTGCTTCCACCCGCTCTCT
GCCCACTCGGGTCTGCAGACCTGGTTCCCACTGAGGCCCTTTGCAAGATGGAATACGGGGCTTACAGGAGCTTTTGTGTGCT
GGTAGAACTATTTCTGTTCAGTCACATGGCATCACTCTTGTACTGCCTGCCACCGCGGAGGAGGCTGGTGACAGGCCAAAG
GCCAGTGAAGAAACACCTTTTCACTCTCAGATCCACTGTGGCACTGGCCACCCCTCCCCAGTACAGGGGTGCTGCAGGTGGCA
GAGTGAATGTCCCCATCATGTGGCCCAACTCTCTGGCTGGCCATCTCCCTCCCAAGAAACAGTGTGCATGGGTTATTTTGG
35 AGTGTAGGTGACTTGTCTTACTCATTTGAAGCAGATTTCTGCTTCTTTTATTTTATAGGAATAGAGGAAGAAATGTCAGATGCG
TGCCAGCTCTTACCCCCCAATCTCTTGGTGGGGAGGGGTGTACCTAAATATTTATCATATCTCTGCCCCCTGAGTGTCTGTTA
GAGAGAAAGAGAACTACTAAGGAAAATAATATTTTAAACTCGCTCCTAGTGTTCCTTTGTGGTCTGTGTACCGTATCTCAG
GAAGTCCAGCCACTTGACTGGCACACACCCCTCCGACATCCAGCGTGACGGAGCCCACTGCCACCTTGTGGCCGCTGAGA
CCCTCGCGCCCCCGCGCCCCCGCGCCCTCTTTTCCCTTGATGGAATTGACCATACAATTTTCACTCTCTTACAGGGAT
40 CAAAAGGACGGAGTGGGGGACAGAGACTCAGATGAGGACAGAGTGGTTTCCAATGTGTTCAATAGATTAGGAGCAGAAATGC
AAGGGGCTGCATGACCTACCAGGACAGAACTTTCCCAATTACAGGGTGACTCACAGCCGATTTGGTGACTCACTTCAATGTGT
CATTTCCGGCTGCTGTGTGTGAGCAGTGGACACGTGAGGGGGGGTGGGTGAGAGAGACAGGCAGCTCGGATTTCAACTACCTTA
GATAATATTTCTGAAAACCTACCAGCCAGAGGGTAGGGCACAAGATGGATGTAATGCACCTTTGGGAGGCCAAGCGGGAGGAT
TGCTTGAGCCAGGAGTTCAAGACCAGCTGGGCAACATACCAAGACCCCGTCTCTTTAAAAATATATATATTTTAAATATAC
45 TTAATATATATTTCTAATATCTTTAAATATATATATATATTTTAAAGACCAATTTATGGGAGAATTGCACACAGATGTGAAAT
GAATGTAATCTAATAGAAGC

237

MLARALLCAVLALSHTANPCCSHPCQNRGVCMSVGFQYKCDCTRTGFGYENCSTPEFLTRIKLFLKPTPNTVHYILTHFKGF
WNVVNNIPFLRNAIMSYVLTSRSHLIDSPPTYNADYGYKSWEAFSNLSYYTRALPPVPDDCPTPLGVKGKKQLPDSNEIVEKLL

CAATTGTCATACGACTTGCAGTGAGCGTCAGGAGCAGCTCCAGGAACCTCTCAGCAGCGCCTCCTTCAGCTCCACAGCCAGACG
CCCTCAGACAGCAAAGCCTACCCCGCGCGCGCCCTGCCCGCGCTCGGATGCTCGCCCGCGCCCTGCTGCTGTGCGCGGTCC
10 TGGCGCTCAGCCATACAGCAAATCCTTGCTGTTCACCCCATGTCAAACCGAGGTGTATGTATGAGTGTGGGATTTGACCAGT
ATAAGTGCATTGTACCCGGACAGGATTCTATGGAGAAAAGTGTCAACACCGGAATTTTGGACAAGAATAAAATTTATTTCTGA
AACCCACTCCAACACAGTGCATACATACTTACCCACTTCAAGGGATTTTGGAAACGTTGTGAATAACATTCCTTCCCTTCGAA
ATGCAATTTATGAGTTATGTCTTGACATCCAGATCACATTTGATTGACAGTCCACCAACTTACAATGCTGACTATGGCTACAAA
GCTGGGAAGCCTTCTCTAACCTCTCCTATTATACCTAGAGCCCTTCCTCCTGTGCTGTGATTTGCCCGACTCCCTTGGGTGTCA
15 AAGGTAAGAAAGCAGCTTCCTGATTCAAAATGAGATTGTGGAAAAATTGCCTTCTAAGAAGAAAGTTTCATCCCTGATCCCCAGGGC
CAAACATGATGTTTTCGATTTCTTTGCCCGACACTTCACGCATCAGTTTTTCAAGACAGATCATAAGCGAGGGCCAGCTTTCACCA
ACGGGCTGGGCCATGGGGTGGACTTAAATCATATTTACGGTGAACACTCTGGCTAGACAGCGTAACTGCGCCTTTTCAAGGATG
GAAAAATGAAATATCAGATAATTGATGGAGAGATGTATCCTCCACAGTCAAAGATACTCAGGCAGAGATGATCTACCCCTCTC
AAGTCCCAGAGCATCAGGTTTGTCTGTGGGCGAGGAGGATTTTGGTCTGGTGGCTGGTGTGATGATGTATGCCACAATCTGGC
20 TGGCGGAACACCAACAGGATGTCGATGTGCTTAAACAGGAGCATCTCGAATGGGGTGTGATGAGCATGTGTTCCAGACAAAGCAGG
TAATACTGATAGGAGAGACTATTAAAGTTGTGATTGAAGATTATGTGCAACACTTGAGTGGCTATCTACTTCAAACGAAATTTG
ACCCAGAACTACTTTTCAACAAACAATTCAGTACCAAAATCGTATTGCTGCTGAATTTAACACCTCTATCTACTGGCATCCCC
TTCTGCTTGACACCTTTTCAAATTCATGACCAGAAATACAACATCAACAGTTTATCTACAACAACCTCTATATTGCTGGAACATG
GAATTACCCAGTTTGTGTAATCATTCACAGGCAAAATGCTGGCAGGGTGTGCTGGTGGTAGGAATGTTCCACCCGAGTACAGA
25 AAGTATCAGAGCTTCCATTGACCAGAGCAGGCAGATGAAATACCAAGTCTTTTATGAGTACCGCAAACGCTTTATGCTGAAGC
CCTATGAATCATTTGAAGAACTTACAGGAGAAAAGGAAATGTCTGCAGAGTTGGAAGCACTCTATGGTGCATCGATGCTGTGG
AGCTGTATCCTGCCCTTCTGGTAGAAAAGCCTCGGCCAGATGCCATCTTTGGTGAACCATGGTAGAAGTTGGAGCACCATTCT
CCTTGAAAGGACTTATGGGTAATGTTATATGTTCTCTGCTACCTGGAAGCCAAGCACTTTTGGTGGAGAAGTGGGTTTTCAAA
TCATCAACACTGCCTCAATTCAAGTCTCTCATCTGCAATAACGTGAAGGGCTGTCCCTTTACTTTCATTAGTGTCCAGATCCAG
30 AGCTCATTTAAACAGTCACCATCAATGCAAGTTCTTCCCGCTCCGGACTAGATGATATCAATCCACAGTACTACTAAAAGAAC
GTTGCACTGAACGTGAGAACTCAATGATCATATTTATTTATTTATATGAACCATGTCTATTAATTTAATTTATTAATAATT
TATATTAAGTCCCTTATGTTACTTAAACATCTTCTGTAACAGAAGTCAGTACTCCTGTTGCGGAGAAAGGAGTCATACTTGTGAA
GACTTTTATGTCTACTCTTAAAGATTTTGTCTGTTGCTGTTAAGTTTGGAAAAAGCTTTTATTTCTGTTTATAAACAGAGAG
AAATGAGTTTTCAGCTCTTTTACTTGAATTTCAACTTATATTTAAGAAGCAAGTAAAGATGTTGAATCTTAAACACTAT
35 CACAAGATGGCAAAATGCTGAAAGTTTTCAGTCTGATGTTTCCAATGCATCTTCCATGATGCATAGAGAAGTAACTAAATGTT
TGAATTTTAAAGTACTTTTGGTTATTTTTCTGTCTATCAACAAAAAGGATATCAGTGCATTATTAATGAATATTTAAATTA
GACATTACCAGTAATTTTCATGTCTACTTTTTTAAATCAGCAATGAAACAATAATTTGAAATTTCTAAATTCATAGGGTAGAATC
ACCTGTAAAGCTTGTGTTGATTCTTAAAGTTATTAACCTGTACATATACCAAAAAAGAGCTGTCTTGGATTTAAATCTGTAA
AATCAGATGAAATTTTACTACAATTGCTTGTGTTAAATATTTTATAAGTGATGTTCCTTTTTACCAAGAGTATAAACCTTTTTA
40 GTGTGACTGTTAAAACTTCCCTTTTAAATCAAAATGCCAAATTTATTAAGTGGTGGAGCCACTGCAGTGTATCTCAAAATAAG
AATATTTTGTGAGATATTCAGAAATTTGTTTATATGGCTGGTAACATGTAAATCTATATCAGCAAAAGGGTCTACCTTTAAA
ATAAGCAATAACAAAGAAGAAAACCAATTTATGTTCAAATTTAGGTTTAAACTTTTGAAGCAAACTTTTTTTTATCCTTGTGC
ACTGCAGGCTGGTACTCAGATTTTGTCTATGAGGTTAATGAAGTACCAAGCTGTGCTTGAATAACGATATGTTTCTCAGATTT
TCTGTTGTACAGTTTAATTTAGCAGTCCATATCACATTTGCAAAAGTAGCAATGACCTCATAAAATACCTCTTCAAAATGCTTAA
45 ATTCATTTTACACATTAATTTTATCTCAGTCTTGAAGCCAATTCAGTAGGTGCAATTGGAATCAAGCCTGGCTACCTGCATGCTG
TTCTTTTCTTTTCTTTTCTTTTAGCCATTTTGCTAAGAGACACAGTCTTCTCATCACTTCGTTTCTCCTATTTTGTTTTACTAGT
TTTAAAGATCAGAGTTCATTTCTTTGGACTCTGCCTATATTTTCTTACCTGAACCTTTTGAAGTTTTCAGGTAACCTCAGCTC
AGGACTGCTATTTAGCTCCTCTTAAAGAAGATTAAGAGAGAAAAAAGGCCCTTTTAAAAATAGTATACACTTATTTTAAAGT
AAAAGCAGAGAATTTTATTTATAGCTAATTTTAGCTATCTGTAACCAAGATGGATGCAAGAGGCTAGTGCCTCAGAGAGAACT
50 GTACGGGGTTGTGACTGGAAGAAAGTTACGTTCCCATCTTAATTAATGAGCCCTTTCTTATTTAAAAACAAACCAATGATATCT
AAGTAGTTCTTCAGCAATAATAATAAGTACGATAATCTTCTTTTTCACATCTCATTTGATCAGTACATGACATTTAATGGTACTGTATA
TCTGAGTCTTTCAGCAATAATAATAAGTACGATAATCTTCTTTTTCACATCTCATTTGATCAGTACATGACATTTAATGGTACTGTATA

340

- TACTTAATTTATTGAAGATTATTATTATGTCTTATTAGGACACTATGGTTATAAACTGTGTTTAAAGCCTACAATCATTGATTT
TTTTTTGTATGTGCACAATCAGTATATTTCTTTGGGGTTACCTCTCTGAATATTATGTAAACAATCCAAAGAAATGATTGTAT
TAAGATTGTGAATAAATTTTAGAAATCTGATTGGCATATTGAGATATTTAAGGTTGAATGTTTGTCTTAGGATAGGCCTAT
GTGCTAGCCCAAAAGAATATTGTCTCATTAGCCTGAATGTGCCATAAGACTGACCTTTTAAAAATGTTTGTAGGGATCTGTGGA
5 TGCTTCGTTAATTTGTTCAGCCACAATTTATTTGAGAAAATATTTCTGTGTCAAGCACTGTGGGTTTAAATATTTTAAATCAAAC
GCTGATTACAGATAATAGTATTTATATAAAATTTGAAAAAATTTCTTTTGGGAAGAGGGAGAAAAATGAAATAAATATCATT
AAAGATAACTCAGGAGAATCTTCTTTACAATTTTACGTTTGAATGTTTAAAGTTAAGAAAGAAATAGTCAATATGCTTGTATA
AAACACTGTTCACTGTTTTTTTTTAAAAAAAACCTTGATTGTTTATTAACATTGATCTGCTGACAAAACCTGGGAATTTGGGTT
GTGTATGCCGAATGTTTCAGTGCCTCAGACAAATGTGTATTTAACTTATGTAAAAGATAAGTCTGGAAATAAATGTCTGTTTATT
10 TTTGTACTATTTA
239
MDPNCSCSCEAGGSCACAGSCKCKKCTSCCKSCCSCPLGCAKCAQGCICKGASEKSCCA
240
CTCCAGTCTCACCTCGGCTTGCAATGGACCCCAACTGCTCCTGCGAGGCTGGTGGCTCCTGCGCCTGCGCCGGCTCCTGCAAGT
15 GCAAAAAGTGCAAAATGCACCTCCTGCAAGAAGAGCTGCTGCTCCTGTTGCCCTTGGGCTGTGCCAAGTGTGCCAGGGCTGCA
TCTGCAAAAGGGCGTCAGAGAAGTGACGCTGCTGTGCTGATGTGCGGACAGCCCTGCTGTGATGAAAAACAGAAATGACACGT
AAAAATCCGAGTGTGTGTCTTCTACAACTCCGACTCATTGCTACATTCTTTTCTGTGAAAATATGTGAATAAATAATTA
CACTTAGACTTGAAAAA
241
20 MDPNCSCSPVGSACAGSCKCKECKCTSCCKSCCSCPVGCAKCAQGCICKGTSDKSCCA
242
CCGCTGCGTGTTCCTCTTGATCGGGAACCTCTGCTTCTCCTTGCTCGAAATGGACCCCAACTGCTCCTGCTCGCCTGTG
CTCCTGTGCTGTGCGGCTCCTGCAAAATGCAAGAGTGCAAAATGCACCTCCTGCAAGAAGAGCTGCTGCTCCTGTGCTGCTGT
GGGCTGTGCAAGTGTGCCAGGGCTGCATCTGCAAGGGAGCTGAGACAAGTGCAGCTGCTGTGCTGATGCCAGGACAGCTG
25 TGCTCTCAGATGTAATAGAGCAACCTATATAAACCTGGATTTTTTTTTTTTTTTTGTACAACCTGACCCGTTGTGATA
TCTTTTTTCTATGAAATATGTGAATGGCAATAAATTCATCTAGACTAAAAA
243
MDPNCSCAAGVSCTCASSCKCKECKCTSCCKSCCSCPVGCAKCAQGCICKGASEKSCCA
244
30 ACTCCGCTTCCACGTGCACCCACTGCCTCTTCCCTTCTCGCTTGGGAACCTAGTCTCGCCTCGGGTTGCAATGGACCCCAAC
TGCTCCTGTGCGCTGGTGTCTCTGCACTGCGCCAGCTCCTGCAAGTGCAAGAGTGCAAAATGCACCTCCTGCAAGAAGAGC
TGCTGCTCCTGCTGCCCTGTGGGCTGTGCCAAGTGTGCCAAGGCTGCATCTGCAAGGGGCATCGGAGAAGTGCAGCTGCTGC
GCCTGATGTGCGGACAGCCCTGCTCCCAAGTACAAATAGAGTGACCCGTAATCTAGGATTTTGTGTTTGTGCTACAATCTT
GACCCCTTGTGCTACATTCCTTTTTTCTGTGAAATATGTGAATAAATAATTAACACTTAG
35 246
TTTAAAGTCAAAATGTTTTATTGTGTCAGTCACATATTTAGTATAAAAAGAAATGCAGCAATGGCTCAGTGTGTATTTTAA
AAAAATCCAGGTGTGTCAGGTGTCTATTACATCTGGGAGAAGAGCTGTTCCACATCAGGCACAGCAGCTGCATTCTCCG
ATGCCCTTTGTCAGACGAGCCCTGGGACACTTGGCACAGCCAGGGGAGCAGGAACAGCAGCTCTTCTTGCAGGAGGTGCATT
TGCATCTTTGCACTTGCAGGAGCCGGCGCACGTGCAGGAGCCA
40 247
MGTQKVTPALIFAITVATIGSFQFGYNTGVINAPEKIIKEFINKTLTDKGNAPPSEVLLTSLWSLSVAIFSVGGMIGSFVGLF
VNRFGRRNSMLIVNLLAVTGGCFMGLCKVAKSVEMLILGRVIGLFCGLCTGFVPMYIGEISPTALRGAFCTLNQLGIVGILV
AQIFGLEFILGSEELWPLLLGFTILPAILQSAALPCFPESPRFLLINRKEENAKQILQRLWGTQDVSQDIQEMKDESARMSQE
KQVTVLELFRVSSYRQPIIIISIVLQLSQQLSGINAVFYSTGIFKDAGVQEPYIATIGAGVNTIFTVWSLFLVERAGRRLHMA

IGLGMAFCSTLMTVSLLLKDNYNMGMSFVCIGAILVFVAFFEIGPGPIPWFIIVAEILFSQGRPAAMAVAGCSNWTSNFLVGLLF
PSAAHYLGAYVFIIFTGFLITFLAFTFFKVPETRGRTFEDITRAFEQAHGADRSGKDGVMEMNSIEPAKETTTNV

248

- GTGGGGTGGGGTGGGGCTGGGGCTTGTGCGCCCTTTTCAGGCTCCACCCTTTGCGGAGATTATAAATAGTCATGATCCCAGCGAG
5 ACCCAGAGATGCCTGTAATGGTGAGACTTTGGATCCTTCTGAGGACGTGGAGAAAACTTTCTGCTGAGAAGGACATTTTGAAG
GTTTTGTGGCTGAAAAAGCTGTTTTCTGGAATCACCCCTAGATCTTTCTTGAAGACTTGAATTAGATTACAGCGATGGGGACAC
AGAAGGTCACCCAGCTCTGATATTTGCCATCACAGTTGCTACAAATCGGCTCTTTCCAATTTGGCTACAAACACTGGGGTCATCA
ATGTCCTCTGAGAAGATCATAAAGGAATTTATCAATAAAACTTTGACGGACAAGGAAATGCCCCACCCTCTGAGGTGCTGCTCA
CGTCTCTCTGGTCTTGTCTGTGGCCATATTTTCCGTGCGGGGTATGATCGGCTCCTTTTCCGTGCGACTCTTCTGTCACCCGCT
10 TTGGCAGGCGCAATTCAATGCTGATTGTCAACCTGTTGGCTGTCACTGGTGGCTGCTTTATGGGACTGTGTAAAGTAGCTAAGT
CGGTTGAAATGCTGATCCTGGGTGCTTGGTTATTTGGCTCTTCTGCGGACTCTGCACAGGTTTGTGGCCATGTACATTGGAG
AGATCTCGCTACTGCCCTGCGGGGTGCCCTTTGGCACTCTCAACCAGCTGGGCATCGTTGTTGGAATTTCTGGTGGCCAGATCT
TTGGTCTGGAATTCATCCTTGGGTCTGAAGAGCTATGGCCGCTGCTACTGGGTTTACCATCCTTCTGCTATCCTACAAAGTG
CAGCCCTTCCATTTTGGCCCTGAAAGTCCAGATTTTGTCTATTAAACAGAAAAGAAGAGGAGAATGCTAAGCAGATCCTCCAGC
15 GGTGTGGGGCACCAGGATGTATCCCAAGACATCCAGGAGATGAAAGATGAGAGTGCAAGGATGTACAAGAAAAGCAAGTCA
CCGTGCTAGAGCTCTTAGAGTGTCCAGCTACCGACAGCCCATCATTTCCATTTGTGCTCCAGCTCTCTCAGCAGCTCTCTG
GGATCAATGCTGTGTTCTATTACTCAACAGGAATCTTCAAGGATGCAGGTGTTCAAGAGCCCATCTATGCCACCATCGGCGCGG
GTGTGGTTAATACATCTTCACTGTAGTTTCTCTATTCTTGGTGAAAGGGCAGGAAGAAGGACTCTGCATATGATAGGCCCTTG
GAGGGATGGCTTTTGTTCACGCTCATGACTGTTCTTGTATTATAAGGATAACTATAATGGGATGAGCTTTGTCTGTATTG
20 GGGCTATCTTGGTCTTTGTAGCCTTCTTTGAAATTTGGACCAGGCCCCATTTCCCTGGTTTATTGTGGCCGAACTCTTCAGCCAGG
GCCCCGCCAGCTGCGATGGCAGTGGCCGGCTGCTCCAACCTGGACCTCCAACCTTCTAGTCCGATTGCTCTTCCCTCCGCTG
CTCACTATTTAGGAGCCTACGTTTATTATCTTCCACCGCTTCTCATTTACCTTCTTGGCTTTTACCTTCTTCAAAGTCCCTG
AGACCCGTGGCAGGACTTTTGAGGATATCACACGGGCTTTGAAGGGCAGGCACACGGTGCAGATAGATCTGGAAGGACGGCG
TCATGGAGATGAACAGCATCGAGCCTGCTAAGGAGACCACCAATGTCTAAGTCGTGCCTCCTTCCACCTCCCTCCCGCAT
25 GGGAAAGCCACCTCTCCCTCAACAAGGAGAGACCTCATCAGATGAACCCAGGACGCTTCTGAATGCTGCTACTTAATTCCTT
TCTCATCCACGCACTCCATGAGCACCCCAAGGCTGCGGTTTGTGGATCTTCAATGGCTTTTAAATTTTATTTCTTGAGCAT
CCTCTTCTGCTTAGGAGAGACCGAGTGAACCTACCTTCATTTTCAAGAGGATTTGGCCGCTTGGCACATGACAACCTTTGCCAGCT
TTTCTCCCTTGGGTTCTGATATTGCCGCACTAGGGGATATAGGAGAGGAAAAGTAAGGTGCAGTTCCCCAACCTCAGACTTA
CCAGGAAGCAGATACATATGAGTGTGGAAGCCGGAGGGTGTATTATGTAAGAGCACCTTCTCACTTCCATACAGCTCTACGTGG
30 CAAATTAACCTGAGTTTATTTATTTATCTCTGGTTTAAATTACATAATTTTTTTTTTTTACTTTAAGTTTCAGGATACATG
TGCCGAATGTGCAGGTTTGTATACATAGGTATATATATGCCATGATGGAATATTTATTTTAAAGCGTAATTTTGCCAAATAA
TAAAAACAGAAGGAAATTGAGATTAGAGGGAGGTGTTTAAAGAGAGGTTATAGACTAGAAGATTGATGCTGGAGAGGTTAAGG
TGCAATAAGAATTTAGGGAGAAATGTTGTTCAATTATTTGGAGGGTAAATGATGTGGTGCCGAGGCTGTACGTTACCTCTTAAC
AATTTCTGTCTTTCAGATGGAACCTTTAACTTCTCGTAAAAGTCATATACCTATATAATAAAGCTACTGATTTCCCTTGAGC
35 TTTTCTCTTAAAGATAATAGTTTACATGTAGTAGTACTTGAATCTAGGATTATTAACATAATATGGGCATGTAGTTAATGATG
GTTGATGGGTTCTAATTTTGGATGGAGTCCAGGGAAGAGAAAGTGATTTCTAGAAAGCCTGTTCCCTCACTGGATGAAATAAC
TCCTTCTGTAGTAGTCTCATTACTTTTGAAGTAATCCCGCCACCTATCTCGTGGGAGAGCCATCCAAATAAGAAACCTAAAAT
AATTTGGTTCTTGGTAGAGATTCAATATTTTCCACTTTGTTCTTTAGGAGATTTTAGGTGTTGATTTTCTGTGTATTTTAACT
CATACCTTTAAAGGAATTTCCCAAAGAATGTTTATAGCAAACTTGGAAATTTGTAACCTCAGCTCTGGGAGAGGATTTTCTCTG
40 AGCGATTATTATCTAAAGTGTGTTGTTGCTTTAGGCTCACGGCACGCTTGGCGTATGCTGTTACCATGTCACTGTGGTCCATG
CCGAATGCCCTCAGGGGACTTGAATCTTTCCAATAAACAGGTTTAGACAGTATGAGTCAATGTGCAGTGTAGCCACACTTGA
GAGGATGAATGTATGTCACTGTCACTTTGCTCTGGGTGGAAGTACGTTATTGTTGACTTATTTCTCTGTGTTTGTTCCTACA
GCCCCCTTTTTCATATGTTGCTCAGTCTCCCTTTCCCTTCTTGGTGCTTACACATCTCAGACCCCTTAGCCAAACCCCTGTCACT
GACAGTATTTTGGTTCTTAGTTCTCACTGTTCCCTCTGCTTCTGGAGCCTTTGAATAAAAAATGCACGTGAGCGCGGATGC
45 GGTGGCTCACGCCCTGTAATCCAGCACTTTGGGAGGCTAGGCGGGCGGTCAGGGGTTTCGAGACCAGTCTGGCCAAACATCGTGA
AACCTGTCTCTACTAAAAATGCAAAAATTAGCCGGGCGTGGTGGCGGGCGCTGTAATCCAGCTACTTGGGAAGCTGAGGCG
GGAGAATCATGTGAACCCGGGACGAGGGGTTGCACTGAGCGGAGATCGCATCATTTGCACTCTAGCCTGGGCCACAGGGCGAGA
CTCCGCTCTCAAAAAAATAAATGCACATAGCTATCGAGTGTGCTTTAGCTTGAAGGAGTGAACCTTGAACCTCATGTCAACTT
TCTGGCTCCTCAACAGTAGGTTGGCAGTAAGGCAAGGCTCCATTTCTCACTGAGAAGATTGTGAATATTTCCATATGGATTTT
50 CTATTGTTACTCTGGTTCTTTGTTTTAAATAAAAAATCTGAATGTACAG

249

MIASHLLAYFFTELNHDQVQKVDQYLYHMRLSDETLLEISKRFKEMEKGLGATTHPTAAVKMLPTFVRSTPDGTEHGEFLALD
LGGTNFRVLWVKVTDNGLQKVEMENQIYAI PEDIMRSGSTQLFDHIAECLANFMDKLQIKDKKLPLGFTTSPCHQTKLDESFL
VSWTKGFKSSGVEGRDVVALIRKAIQRRGDFDIDIVAVVNDTVGTMTCGYDDHNCIEGLIVGTGSNACYMEEMRHIDMVEGDE
5 GRMCINMEWGAFGDDGSLNDRTEFDQIEDMSLNPQKQLEFKMISGMYMGELVRLILVKMAKEELLFGGKLSPELLNTGRFET
KDISDIEGEKDGIRKAREVLMRLGLDPTQEDCVATHRICQIVSTRSASLCAATLA AVLQRIKENKGEERLRSTIGVDGSVYKXH
PHFAKRLHKTVRRLLVPGCDVRFLRSEDGSGKGAAMVTAVAYRLADQHRARQKTLEHLQLSHDQLLEVKRRMKVEMERGLSKETH
ASAPVKMLPTTYVCATPDGTEKDFLALDLGCTNFRVLLVRVRNGKWGGVEMHNKIYAI PQEVMHGTGDELFDHIVQCIADFLEY
MGMKGVSLLPLGFTTSPPCQNSLDESILLKWTGFKASGCEGEDVVTLLEKAIHRREEFDLDVVAVVNDTVGTMTCGFEDPHC
10 EVGLIVGTGSNACYMEEMRNVELVEGEEGRMCVNMEWGAFGDNGCLDDFRTEFDVAVDEL SLNPGKQRFKMI SGMVYLGEIVRN
ILIDFTKRGLLFRGRISERLKTGRIFETKFLSQIESDCLALLQVRAILQHLGLESTCDDSIIVKEVCTVVARRAAQLCGAGMAA
VVDRIRENRGLDALKVTGVDGTLTKLHPHFAKVMHETVKDLAPKCDVSFLQSEDGSGKGAALITAVACRIREAGQR

250

GTTGCATGAAACTCCGGCGCAGGAGTCCCGGGCTGCCGCTGGCAACATCGTGTCACCCAGCTAAGAAAATCCGGGGCCCCGAGC
15 CACGCGCCTGTGAATCGGAGAGGTCCCACTGCCCCGAGTGGAGCCGGGCTGAGATTCTTCTCAAGTTGAGCCTCAGTGATCCTGT
GGCCGAAGTTAGCGCCTTGACGTGGGACAACCGGACACGTCCGACGAGAGAAGTACGAGCGCCTTCTAGCAGTTGTGACGCCAA
AATACAGTCTCCGGAGACCCGCGCCTCCGCGAGCCGGGCGCACCTCGCCGGTAGCCTTCTTTGTGCGCCGTCCGGACTCCCA
GCTCCCGGGCCCGGAGCCGAGCCCGAGCCAGCACAAGCAGTCCGACCGCGCCCGCCGCTCCCTCTCGCGTCTCCGCTCGGTTTC
CCAATCTGCGCCGTCGGGCGCGGCGAGGATGATTGCTCGCATCTGCTTGCTTCTTACCGAGCTCAACCATGACCAAG
20 TGCAGAAGGTTGACCAATCTCTACCATGCGCCTCTCTGATGAGACCTCTTGGAGATCTCTAAGCGGTTCCGCAAGGAGA
TGGAGAAAGGGCTTGGAGCCACCACTCACCTACTGCAGCAGTGAAGATGCTGCCACCTTTGTGAGGTCCACTCCAGATGGGA
CAGAACACGGAGAGTTCTTGGCTCTGGATCTTGGAGGGACCAACTTCCGTGTGCTTTGGGTGAAAGTAACGGACAATGGGCTCC
AGAAGGTGGAGATGGAGAATCAGATCTATGCCATCCCTGAGGACATCATGCGAGGCAGTGGCACCCAGCTGTTTGACCACATTG
CCGAATGCCTGGCTAACTTCAATGGATAAGCTACAAATCAAGACAAGAAGCTCCCACTGGGTTTACCTTCTCGTTCCCTTGCC
25 ACCAGACTAAACTAGACGAGAGTTTCTTGGTCTCATGGACCAAGGAGTTCAAGTCCAGTGGAGTGGGAAGGCAGAGAGCTTGTGG
CTCTGATCCGGAAGGCCATCCAGAGGAGAGGGGACTTTGATATCGACATTGTGGCTGTGGTGAATGACACATGGAGAGATGCA
TGACCTGTGGTTATGATGACCACAACCTGTGAGATTGGTCTCATTTGTGGGCACGGGCAGCAACGCTGCTACATGGAGAGATGC
GCCACATCGACATGGTGGAAAGGCATGAGGGGCGGATGTGTATCAATATGGAGTGGGGGGCCTTCGGGGACGATGGCTCGCTCA
ACGACATTCGCACTGAGTTTGACAGGAGATTGACATGGGCTCACTGAACCCGGGAAAGCAACTGTTTGAGAAGATGATCAGTG
30 GGATGTACATGGGGGAGCTGGTGAGGCTTATCTTGTGAAGATGGCCAAGGAGGAGCTGCTCTTTGGGGGGAAGCTCAGCCCAG
AGCTTCTCAACACCGGTGCTTTGAGACCAAGACATCTCAGACATTGAAGGGGAGAAGGATGGCATCCGGAAGGCCCGTGAGG
TCCTGATGCGGTGGGCTTGACCCGACTCAGGAGGACTGCGTGGCCACTACCGGATCTGCCAGATCGTGCCACACGCTCCG
CCAGCTGTGCGCAGCCACCTGGCCGCGGTGCTGCAGCGCATCAAGGAGAACAAGGCGAGGAGCGGCTGCGCTCTACTATTG
GGGTGACGCTTCCGTCTACAAGAAACACCCCATTTTGCCAAAGCTCTACATAAGACCGTGCGGCGGCTGGTGCCCGGCTGCG
35 ATGTCCGCTTCCCTCGCTCCGAGGATGGCAGTGGCAAGGTGACAGCATGGTGACAGCAGTGGCTTACCGCTGGCCGATCAAC
ACCGTGCCCGCCAGAAGACATTAGAGCATCTGCAGCTGAGCCATGACCAGCTGCTGGAGGTCAAGAGGAGGATGAAGGTAGAAA
TGGAGCGAGGTCTGAGCAAGGAGACTCATGCCAGTGCCCCGCTCAAGATGCTGCCACCTACGTGTGTGTACCCCGGACGGCA
CAGAGAAAGGGGACTTCTTGGCCTTGACCTTGAGGAACAATTTCCGGGTCTGCTGGTCCGCTGTCGGAATGGGAAGTGGG
GTGGAGTGGAGATGCAACAAGATCTACGCCATCCCGCAGGAGGTCATGCACGGCACCGGGGACGAGCTCTTTGACCACATTG
40 TCCAGTGCATCGCGGACTTCTCGAGTACATGGGCATGAAGGGCGTGTCCCTGCCCTCTGGGTTTACCTTCTCCTTCCCTTGCC
AGCAGAACAGCCTGGACGAGAGCATCTCTCAAGTGGACAAAAGGCTTCAAGGCATCTGGCTGCGAGGGCGAGGACGTGGTGA
CCCTGCTGAAGGAAGCGATCCACCGCGAGAGGAGTTTGACCTGGATGTGGTTGCTGTGGTGAACGACACAGTCGGAACATATGA
TGACCTGTGGCTTTGAAGACCTCACTGTGAAGTTGGCCTCATTTGTGGCACGGGCAGCAATGCCCTGCTACATGGAGGAGATGC
GCAACGTGGAACTGTGGAAGGAGAAGAGGGCGGATGTGTGTAACATGGAATGGGGGGCATTCGGGGACAAATGGATGCTTAG
45 ATGACTTCCGCACAGAATTTGATGTGGCTGTGGATGAGCTTTCACCAACCCCGGCAAGCAGAGGTTCCGAGAAAATGATCAGTG
GAATGTACCTGGGTGAGATTGTCCGTAACATTCTCATCGATTTACCAAGCGTGGACTGCTCTTCCGAGGCCGATCTCAGAGC
GGCTCAAGACAAGGGGCATCTTTGAAACCAAGTTCTTGTCTCAGATTGAGAGTGACTGCTTGGCCCTGCTGCAAGTCCGAGCCA
TCCTGCAACACTTAGGGCTTGAGAGCACCTGTGACGACAGCATCAITGTTAAGGAGGTGTGCACTGTGGTGGCCCGCGGGCAG
CCCAGCTCTGTGGCGCAGGCATGGCCGCTGTGGTGGACAGGATACGAGAAAACCGTGGGCTGGACGCTCTCAAGTGACAGTGG
50 GTGTGGATGGGACCTCTACAAGCTACATCTCCTCACTTTGCCAAAGTCATGCATGAGACAGTGAAGGACCTGGCTCCGAAATGTG
ATGTGTCTTCTTCTGACGTGAGGATGGCAGCGGGAAGGGGCGGCGCTCATCACTGCTGTGGCTGCCGATCCGTGAGGCTG

GACAGCGATAGAACCCCTGAAATCGGAAGGGACTTCCTCTTTCTCTCCTTCCTCCCTGTTTTAAATTATAAGATGTCATCCCCCT
TGTCAGAGACAGACCCCTTGGCTTTTGCTTGGCAGAGAGGACCCCACTGGACTGGGTTTTGTCTCTGCATCTCATTGTAGAG
CTTGGTGGCTGAGCTTGGCCCTATTAAAGATAAATAGAGTTCCAAATAAGGATTTGTTACATGCATCATAACCATTTCCATTGG
TTCTCCTAAAACATGAAAATTATCCTTCTAGTAATCCCCCTTGCCAAATCCATGTCCCTGTATAATTCTACAGGATGGGGAC
5 ACTAATGAAGATACGGTTGCTTCACCTTGGAGCCTGAACATGACATTTCTAAGTGGGGTGCATCCCCCAGCACATGATGTTGTTA
CTGATTCCTCTGTCAGAGATCTGGGAGGTCTCCACTGAGGATGTGAGCCTGATTATCCTATAGGCAGACGTGGGGAGGGTGGAG
GGGTGACAGTGGAGGAAAAATCCATGGATATCCACGCAGCAGCCCTCTTTAACCTCATCTACAAGCATTGCGCTGTGGATTCC
AGCATTGCGCATTCCTGGAATCAAGGAATCCTGAGTCTGGGCAATGAAACCAAGCCAGGAGTTGACGCATCCTGCAGTTGGGC
CAGCTGTGCGATCTCAGCGGGGCGCACATGTTATCCACAAGCAATGGACCTTTGGGGAAGGGGAGTTTTAGTTTGTTTTACA
10 AATTTTTCTGCAAAAGTGAATCACTGTATTTTCATTTAATTTATATTTGAAATTTTATTTAGTTCTTGAGTAGATCTGCTT
CTTCATCTTGACATGTAATGAATGGTCAGTTGTACGTAATGTATTTATATGTTAATTTGTTATGTATATAGATGTGCAAGTCTT
GTCAGAATTGGCCTCAGTGTAGTTAAAGGGCAGAAGGGGAAGATACTGACTAGTCATAGAAATACCTCATTCGCTGTGGGAAG
AGAAGGGAAGCCTCTTCAGGGTGAATGGCAAGCGGTTGCTTCTCCG

251

15 MTSKLAVALLAFLISAALCEGAVLPRSAKELRCQCIKTYSPKPFHPKFIKELRVIESGPHCANTEIIVKLSDGRELCLDPKENW
VQRVVEKFLKRAENS

252

AGCAGAGCACACAAGCTTCTAGGACAAGAGCCAGGAAGAAACCACCGGAAGGAACCATCTCACTGTGTGTAACATGACTTCCA
AGCTGGCCGTGGCTCTCTTGGCAGCCTTCTGATTTCTGCAGCTCTGTGTGAAGGTGCAGTTTGGCCAAGGAGTGCTAAAGAAC
20 TTAGATGTCAAGTGCATAAAGACATACTCCAAACCTTTCCACCCCAAATTTATCAAAGAACTGAGAGTGATTGAGAGTGGACCAC
ACTGCGCCAAACACAGAAATATTGTAAAGCTTTCTGATGGAAGAGAGCTCTGTCTGGACCCCAAGGAAACTGGGTGCAGAGGG
TTGTGGAGAAGTTTTTGAAGAGGGCTGAGAATTCATAAAAAAATTCATTCTCTGTGGTATCCAAGAATCAGTGAAGATGCCAGT
GAACTTCAAGCAAATCTACTTCAACACTTCATGTATTGTGTGGGTCTGTTGTAGGGTTGGCAGATGCAATACAAGATTCTCGG
TTAAATTTGAATTTTCAGTAAACAATGAATAGTTTTCATTTGTACCATGAAATATCCAGAACATACTTATATGTAAAGTATTATF
25 TATTTGAATCTACAAAAACAACAAATAATTTTTAAATATAAGGATTTTCTTAGATATTGCACGGGAGAATATACAAATAGCAA
AATTGAGCCAAGGGCCAAGAGAATATCCGAACCTTAATTTTCAGGAATGAATGGGTTTGTCTAGAAATGTGATTTTGAAGCATCA
CATAAAATGATGGGACAATAAATTTTGCATAAAGTCAAATTTAGCTGGAAATCCTGGATTTTTTCTGTTAAATCTGGCAAC
CCTAGTCTGCTAGCCAGGATCCACAAGTCTTGTTCCTGCTGCTTGGTTTCTCCTTTATTTCTAAGTGGAAAAAGTATTAGC
CACCATCTTACCTCACAGTGAATGTTGTGAGGACATGTGGAGCACTTTAAGTTTTTTCATCATAACATAAATTTTCAAGTG
30 TAACTTATTAACTATTATTTATTTATGTTATTTAAGCATCAAATTTGTGCAAGAATTTGGAAAAATAGAAGATGAATC
ATTGATTGAATAGTTATAAAGATGTTATAGTAAATTTATTTTATTTTAGATATTAAATGATGTTTTATTAGATAAATTTCAATC
AGGGTTTTTAGATTAAACAAGAAACAATTTGGGTACCCAGTTAAATTTTCATTTTCAGATAAAACAACAATAATTTTTTAGTATA
AGTACATTATTGTTTATCTGAAAGTTTAAATGAACTAACAAATCCTAGTTTGATACTCCAGTCTTGTCATTGCCAGCTGTGTT
GGTAGTGTGTTGAATTACGGAATAATGAGTTAGAATTTAAACAGCCAAAATCCACAGTCAATATTAGTAATTTCTTG
35 CTGGTTGAAACTTGTATTATGTTACAAATAGATTCTTATAATATTATTTAAATGACTGCATTTTAAATACAAGGCTTTATAT
TTTTAACTTTAAGATGTTTTTATGTGCTCTCCAAATTTTTTTTACTGTTTCTGATTGTATGGAATATAAAAGTAAATATGAAA
CATTTAAATATAAATTTGTTGTCAAAGTAAAAAATAAATAA

253

MGKVKVGVNGFGRIGRLVTRAAFNSGKVDIVAINDPFIDLNYMVYMFQYDSTHGKPHGTVKAENGKLVINGNPITIFQERDPSK
40 IKWGDAGAEYVVESTGVFTTMEKAGAHLQGGAKRVII SAPSADAPMFVMGVNHEKYDNLKII SNASCTTNCLAPLAKVIHDF
GIVEGLMTTVHAITATQKTVDPGSGKLWRDGRGALQNIIPASTGAAKAVGKVIPELNGKLTGMAFRVPTANVSVDLTCRLEKP
AKYDDIKKVVQASEGPLKGILGYTEHQVSSDFNSDTHSSTFDAGAGIALNDHFVKLISWYDNEFGYSNRVVDLMAHMASKE

254

CTCTCTGCTCCTCCTGTTTCGACAGTCAGCCGCATCTTCTTTTGGCTCGCCAGCCGAGCCACATCGCTCAGACACCATGGGGAAG
45 GTGAAGGTGCGAGTCAACGGATTGTCGTATTGGGCGCCTGGTCACCAGGGCTGCTTTTAACTCTGGTAAAGTGGATATTGTT
GCCATCAATGACCCCTTCATTGACCTCAACTACATGGTTTACATGTTTCCAAATATGATTCCACCCATGGCAAATTCATGGCACC
GTCAAGGCTGAGAACGGGAAGCTTGTCATCAATGGAATCCCATCACCATCTTCCAGGAGCGAGATCCCTCCAAATCAAGTGG
GGCGATGCTGGCGCTGAGTACGTCTGGAGTCCACTGGCGTCTTACCACCATGGAGAAGGCTGGGCTCATTTGCAGGGGGGA

GCCTAAAGGGTCATCATCTCTGCCCCCTCTGCTGATGCCCCCATGTTTCGTTCATGGGTGTGAACCATGAGAAGTATGACAACAGC
CTCAAGATCATCAGCAATGCCTCCTGCACCACCAACTGCTTAGCACCCCTGGCCAAGGTCATCCATGACAACCTTTGGTATCGTG
GAAGGACTCATGACCACAGTCCATGCCATCACTGCCACCAGAACTGTGGATGGCCCCCTCCGGGAACTGTGGCGTGATGGC
CGCGGGGCTCTCAGAACATCATCCCTGCCTCTACTGGCGCTGCCAAGGCTGTGGGCAAGGTCATCCCTGAGCTGAACGGGAAG
5 CTCACTGGCATGGCCTTCCGTGTCCCCACTGCCAACGTGTCAGTGGTGGACCTGACCTGCCGTCTAGAAAACTGCCAAATAT
GATGACATCAAGAAGGTGGTGAAGCAGGCGTCGGAGGGCCCCCTCAAGGGCATCCTGGGCTACACTGAGCACCAGGTGGTCTCC
TCTGACTTCAACAGCGACACCCACTCCTCCACCTTTGACGCTGGGGCTGGCATTGCCCTCAACGACCACTTTGTCAAGCTCAT
TCCTGGTATGACAACGAATTTGGCTACAGCAACAGGGTGGTGGACCTCATGGCCCATGAGGCTCAAGGAGTAAGACCCCTGG
ACCAACAGCCCCAGCAAGAGCACAAGAGGAAGAGAGACCCCTCACTGCTGGGAGTCCCTGCCACACTCAGTCCCCCACCACA
10 CTGAATCTCCCCCTCCTCAGTTGCCATGTAGACCCCTTGAAGAGGGGAGGGGCTAGGGAGCCGACCTTGTCTATGTACCATC
AATAAAGTACCTGTGCTCAACC
255
MSLSNKLTLDKLDVKGKRVVMRVDNFNPMKNNQITNNQRIKAAVPSIKFCLDNGAKSVVLSHSLGRPDGVMPDKYSLEPVAVE
LKSLLGKDVFLKDCVGPVEKACANPAAGSVILLENLRFHVEEKGKGDASGNKVKAEPKIEAFRASLSKLDVYVNDVAFGT
15 AHRHSSMVGVLNPKAGGFLMKKELNYFAKALESPERPFLLAILGGAKVADKIQILNNMLDKVNEMIIGGGMAFTFLKVLNNME
IGTSLFDEEGAKIVKDLMSKAEKNGVKITLPVDFVTADKFDENAKTGQATVASGIPAGWMLDCGPESKKYAEAVTRAKQIVW
NGPVGVFWEAFARGTKALMDEVVKATSRGCITIIIGGGDTATCCAKWNTEDKVS HVSTGGGASLELLEKVLPGVDALSNI
256
AAGCCTCCGGAGCGCAGCTCGGCAGTCGGCTCCCTCGTTGACCGAATCACCGACCTCTCTCCCCAGCTGTATTTCCAAAATGTC
20 GCTTTCTAACAAGCTGACGCTGGACAAGCTGGACGTTAAAGGGAAGCGGGTCGTTATGAGAGTCCGACTTCAATGTCTCCATGAA
GAACAACCAGATAACAACAACCAGAGGATTAAGGCTGCTGTCCCAAGCATCAAATTCTGCTTGGACAATGGAGCCAAGTCGGT
AGTCCTTATGAGCCACCTAGGCGCGCCTGATGGTGTGCCCATGCCTGACAAGTACTCCTTAGAGCCAGTTGCTGTAGAAGTCAA
ATCTCTGCTGGGCAAGGATGTTCTGTTCTGAAGGACTGTGTAGGCCCAGAAGTGGAGAAAGCCTGTGCCAACCCAGCTGCTGG
GTCTGTCTCCTGCTGGAGAACCCTCCGCTTTCATGTGGAGGAAGAAGGGAAGGAAAGATGCTTCTGGGAACAAGGTTAAAGC
25 CGAGCCAGCCAAAATAGAAGCTTTCCGAGCTTCACTTTCCAAGCTAGGGGATGTCTATGTCAATGATGCTTTTGGCACTGCTCA
CAGAGCCCACAGCTCCATGGTAGGAGTCAATCTGCCACAGAAGGCTGGTGGGTTTGTGATGAAGAAGGAGCTGAAGTACTTTG
AAAGGCTTGGAGAGCCACAGAGCGACCTTCTTGGCCATCTTGGCGGAGCTAAAGTTGCAGACAAGATCCAGCTCATCAATAA
TATGCTGGACAAAGTCAATGAGATGATTATTTGGTGGTGGAAATGGCTTTTACCTTCTTAAAGTGTCAACAACATGGAGATTGG
CACTTCTCTGTTTGTGAAGAGGGAGCCAAAGATTGTCAAGACCTAATGTCCAAGCTGAGAAGAATGGTGTGAAGATTACCTT
30 GCCTGTTGACTTTGTCACTGCTGACAAGTTTGTAGAGAATGCCAAGACTGGCCAAGCCACTGTGGCTTCTGGCATACTGCTGG
CTGGATGGGCTTGGACTGTGGTCTGAAAGCAGCAAGAAGTATGCTGAGGCTGTCACTCGGGCTAAGCAGATTGTGTGGAATGG
TCCTGTGGGGTATTTGAATGGGAAGCTTTTGGCCGGGGAACCAAGCTCTCATGGATGAGGTGGTGAAGCCACTTCTAGGGG
CTGCATCACCATCATAGGTGGTGGAGACTGCCACTTGTGTGCCAAATGGAACACGGAGGATAAAGTCAGCCATGTGAGCAC
TGGGGGTGGTGCCAGTTTGGAGCTCCTGGAAGGTAAAGTCTTCTGCGGTGGATGCTCTCAGCAATATTTAGTACTTTCTGTC
35 CTTTGTAGTTCTGTGCACAGCCCCTAAGTCAACTTAGCATTTTCTGCATCTCCACTTGGCATTAGCTAAACCTTCCATGTCAA
GATTGAGCTAGTGGCCAAGAGATGCAGTGCAGGAACCCCTAAACAGTTGCACAGCATCTCAGCTCATCTTCACTGCACCCCTGG
ATTTGCATACATTTCTCAAGATCCCATTGAATTTTGTAGTACTAAACCATTTGTGATTTCTAGAGTGCATATATTTATATTTT
GCCTGTAAAAAGAAAGTGAAGCAGTGTAGCTTAGTCTCTTTGTAGTATAGTTATGATTAGCTTTGTCACTGTTTCACTA
CTCAGCATGGAACAAGATGAAATTCATTTGTAGGTAGTGAGACAAAATTGATGATCCATTAAGTAAACAATAAAGTGTCCA
40 TTG
257
MSILKIHAREIFDSRGNPTVEVDLFTSKGLFRAAVPSGASTGIYEALRLDNDKTRYMGKGVSKAVEHINKTIAPALVSKKLNV
TEQEKIDKLMIEMDGTENKSKFGANAILGVSLAVCKAGAVEKGVPLRYHIADLAGNSEVILPVPAPNVINGGSHAGNKLAMQEF
MILPVGAAANFREAMRIGAEVYHNLKNVKEKYGKDNTNVDGEGGAPNILENKEGLELLKTAIGKAGYTDKVVIGMDVAASEFF
45 RSGKYDLDFKSPDDPSRYISPDQLADLYKSFIDYPVVSIEDPFDQDDWGAQKFTASAGIQVVGDDLTFTNPKRIAKAVNEKS
CNCLLLKVNQIGSVTESLQACKLAQANGWGMVSHRSGETEDTFIADLVVGLCTGQIKTGAPCRSERLAKYNQLLRIEELGSK
AKFAGRNFRNPLAK
258

25 MPHSYPALSAEQKKELSDIALRIVAPGKGILAADESVGSMAKRLSQIGVENTEENRRLRYQVLF SADDRVKKCIGGVIFPHETL
YQKDDNGVPFVRTIQDKGIVGIGKVDKGVPLAGTDGETTTQGLDGLSERCAQYKKDGADFAKWRCVLKISERTPSALAILENA
NVLARYASICQQNGIVPIVEPEILPDGDHDLKRCQYVTEKVLAAVYKALSDHHVYLEGTLTKPNMVTPGHACPIKYTPPEETAMA
TVTALRRTVPPAVPGVTFLSGGQSEAEASFNLNAINRCLPRPWALTFYSGRALQASALNAWRGQRDNAGAATEEFIKRAEVNG
LAAOGKYGSGEDGGAAOSLYIANHAY

[illegible]

261

MAPSRKFFVGGNWKMNKRKQSLGELIGTLNAAKVPADTEVVCAPPTAYIDFARQKLDPKIAVAAQNCYKVTNGAFTGEISPGMI
KDCGATWVVLGHSERRHVFGESEDELIGQKVAHALAEBGLGVIACIGEKLDEREAGITEKVVPEQTKVIADNVKDWKVVLAYEPV
WAIGTGKTATPQQAQEVHEKLRGWLKSNVSDAVAQSTRIYGGSVTGATCKELASQPDVDGFLVGGASLKPEFVDIINAKQ

262

- 5 GCGCGACACTGACCTTCAGCGCCTCGGCTCGGCCATGGCGCCCTCCAGGAAGTTCTTCGTTGGGGGAAACTGGAAGATGAACGG
GCGGAAGCAGAGTCTGGGGGAGCTCATCGGCACTCTGAACGCGGCCAAGGTGCCGCGCCACACCGAGGTGGTTTGTGCTCCCCC
TACTGCCTATATCGACTTCGCCCCGAGAGCTAGATCCCAAGATTGCTGTGGCTGCCGAGAACTGCTACAAAGTGACTAATGG
GGCTTTTACTGGGGAGATCAGCCCCTGGCATGATCAAAGACTGCCGAGCCACGTGGGTGGTCTGGGACACTCAGAGAGAAGGCA
TGCTTTTGGGGAGTCAGATGAGCTGATTGGGCAGAAAGTGGCCCATGCTCTGGCAGAGGGACTCGGAGTAATCGCCTGCATTTGG
10 GGAGAAGCTAGATGAAAGGGAAGCTGGCATCACTGAGAAGGTTGTTTTCGAGCAGACAAAGGTCATCGCAGATAACGTGAAGGA
CTGGAGCAAGGTCTCTCTCGCCTATGAGCCTGTGTGGGCCATTGGTACTGGCAAGACTGCAACACCCCAACAGGCCAGGAAGT
ACACGAGAAGCTCCGAGGATGGCTGAAGTCCAACGTCTCTGATGCGGTGGCTCAGAGCACCCGTATCATTTATGGAGGCTCTGT
GACTGGGGCAACCTGCAAGGAGCTGGCCAGCCAGCCTGATGTGGATGGCTTCCTTGTGGGTGGTGTTCCTCAAGCCCGAATT
CGTGGACATCATCAATGCCAAACAATGAGCCCCATCCATCTTCCTTACCCTTCCTGCCAAGCCAGGGACTAAGCAGCCAGAAAG
15 CCCAGTAACGCCCCCTTCCTGCATATGCTTCTGATGGTGTATCTGCTCCTTCCTGTGGCCTCATCCAACTGTATCTTCCTT
TACTGTTTATATCTTCACCTGTAAATGGTTGGGACCAGGCCAATCCCTTCTCCACTTACTATAATGGTTGGAACATAACGTCAC
CAAGGTGGCTTCTCCTTGGCTGAGAGATGGAAGCGGTGGTGGGATTGCTCCTGGGTTCCTTAGGCCCTAGTGAGGGCAGAAGA
GAAACCATCTCTCCCTTCTTACACCGTGAGGCCAAGATCCCTCAGAAGGCAGGAGTGTGCCCTCTCCATGGTGGCCGTGC
CTCTGTGCTGTGTATGTGAACACCCATGTGAGGGAATAAACCTAGCACTAGG

20 263

MASKLLRAVILGPPGSGKGTVCQRIQNFLQLHSSGHFLRENIKASTEVGEMAKQYIEKSLVDPDHVITRLMSELENRRGQH
WLLDGFPRTLGQAEALDKICEVDLVISLNI PFETLKDRLSRRIWHPPSGRVYNLDFNPPHVHGIDDVTEPLVQEQEDDKPEAVA
ARLRQYKDVAKPVIELYKSRGVLHQPSGTETNKIWPVYTLFSNKITPIQSKEAY

264

- 25 CGGCGCTGGGTGAGGGGAGGGGTTGTCTTAAAGTCTCTCCTTCCCCCTGTAGGGGCGGCCGCGAGTCCCAGTGAGAGCGGA
GGGTGCCAGAGGTAGGGGGCGAGAAACAAAGTTCCCGGGCTTCCCTCCGGGGCCGCGGTGGGGCTGCGCGTTTGACCGCCCC
CCTCCTCGCGAAGCAATGGCTTCCAAACTCTGCGCGCGGTATCTCCTCGGGCCGCCCGGCTCGGGCAAGGGCACCGGTGTGCCAG
AGGATCGCCAGAACTTTGGTCTCCAGCATCTCTCCAGCGGCCACTTCTTGGGGAGAACATCAAGGCCAGCACCGAAGTTGGT
GAGATGGCAAGCAGTATATAGAGAAAAGTCTTTTGGTTCCAGACCATGTGATCACACGCCTAATGATGTCCGAGTTGGAGAAC
30 AGGCGTGGACAGCACTGGCTCCTTGATGGTTTCTTAGGACATTAGGACAAGCCGAAGCCCTGGACAAAATCTGTGAAGTGGAT
CTAGTGATCAGTTTGAATATTCCATTTGAAACACTTAAAGATCGTCTCAGCCGCGTTGGATTACCCCTCCTAGCGGAAGGGTA
TATAACCTGGACTTCAATCCACCTCATGTACATGGTATTGATGACGTCACTGGTGAACCGTTAGTCCAGCAGGAGGATGATAAA
CCCGAAGCAGTTGCTGCCAGGCTAAGACAGTACAAAGACGTGGCAAAGCCAGTCATTGAATTATACAAGAGCCGAGGAGTGCTC
CACCAATTTCCGGAACGGAGACGAACAAAATCTGGCCCTACGTTTACACACTTTTCTCAAACAAGATCACACCTATTTCAGTCC
35 AAAGAAGCATATTGACCTGCCCAATGGAAGAACCGGAAGATGTGGTCATTTCATTCAATAGTGTGTAGTATTGGTGTCTGTG
TCCAAATTAGAAGCTAGCTGAGGTAGCTTGACAGCATCTTTTCTAGTTGAAATGGTGAACATGATAGGAAAACAAATGAGTAGAAA
GAGTTTCATGAAGAGGCCCTCCTCTGCCTTTCAAAGGCTGGTCACCTACACATGTTTAAAGGTGTCTCTGCACATGTCTCAAGCC
CATCACAAGAAAGCAAGTACAGTGTGGATTTCAAATGGTGTGTAACCTCAGCTCCAGCTGGTTTTTGCAGCTGTTGTCTGTGGT
AATAATTTTGCATGTGATGGTGATAGTCTCTGGTTCTCCCATCCCCACAAAGGCTGTTGAACCACAGCACCAGGAAGCCTGA
40 GAATGAATCCTGAGGGCTCTAGCCCAGGCTTTGTCCCAGGCTTCTGGTGTGTGCCCTCCTGGTAAACAGTGAAATTGAAGCTAC
TTACTCATAGTGGTTGTTTCTCTGGTCTTGAGTGACTGTGTCCACAGTTCAATTTTTTCCGGTAGGAATAACTCCTTTTCTACA
TCCACGCTCCATAGAGTCTCTCCTTTTCAGACATCTGGGATGAAAGAATTTGGCTTTTTTTTTTCTTTTTTTTGGACATCT
GTTTTCACTCTTAGGCTTTTAAACAATAGTTATTGCTTTTATCCCTCAGATTCTAATAACTGAGAGCGATGGGGCTATATTG
AATCTCTGTATGCACTGAGAACTGAGCTATGAAGAGAATCTTATTAAACTGCTGGTCTGACTTTATGGATTGACACTGTTTCCTT
45 TCTTTTATTGTGAAAAAAAAAAAAAAAA

265

MDPNCSCAAGDSCTCAGSCKCKECKCTSCCKSCCSCCPVCGAKCAQGCICKGASDKCSCCA

266

ATGGATCCCAACTGCTCCTGCGCCGCCGGTGACTCCTGCACCTGCGCCGGCTCCTGCAAAATGCAAAGAGTGCAAAATGCACCTCC
TGCAAGAAAAGCTGCTGCTCCTGCTGCCCCGTGGGCTGTGCCAAGTGTGCCCAGGGCTGCATCTGCAAAGGGGCGTCGGACAAG
TGCAGCTGCTGCGCCTGA

5 267

MRIAVICFCLLGITCAIPVKQADSGSSEKQLYNKYPDVATWLNPDPSQKQNLAPQTLPSKSNESHDMDDMEDDDDDHVD
SQDSIDSNDSDVDVDDTDDSHQSDSHHSDESEDELVTDFPTDLPAVEVFTPVVPTVDTYDGRGDSVVYGLRSKSKKFRRPDIQYP
DATDEDITSHMESEELNGAYKAIPVAQDLNAPSDWDSRGKDSYETSQDLDQSAETHSHKQSRLYKRKANDESNEHSDVIDSQEL
SKVSREFHSHEFHSHEDMLVVDPKSKEEDKHLKFRISHELDSASSEVN

10 268

GCAGAGCACAGCATCGTCGGGACCAGACTCGTCTCAGGCCAGTTGCAGCCTTCTCAGCCAAACGCCGACCAAGGAAAACACTCACT
ACCATGAGAATTGCAGTGATTTGCTTTTGCCCTCCTAGGCATCACCTGTGCCATACCAGTTAAACAGGCTGATTCTGGAAGTTCT
GAGGAAAAGCAGCTTTACAACAAATACCCAGATGCTGTGGCCACATGGCTAAACCCTGACCCATCTCAGAAGCAGAATCTCCTA
GCCCCACAGACCTTTCCAAGTAAGTCCAACGAAAGCCATGACCACATGGATGATATGGATGATGAAGATGATGATGACCATGTG
15 GACAGCCAGGACTCCATTGACTCGAAGACTCTGATGATGTAGATGACACTGATGATTCTCACCAGTCTGATGAGTCTCACCAT
TCTGATGAATCTGATGAACGGTCACTGATTTTCCACGGACCTGCCAGCAACCGAAGTTTCTACTCCAGTTGTCCCCACAGTA
GACACATATGATGGCCGAGGTGATAGTGTGGTTTATGGACTGAGGTCAAAATCTAAGAAGTTTCGAGAGCTGACATCCAGTAC
CCTGATGCTACAGACGAGGACATCACCTCACACATGGAAGCGAGGAGTTGAATGGTGCATACAAGGCCATCCCCGTGGCCAG
GACCTGAACCGCCTTCTGATTTGGGACAGCCGTGGGAAGGACAGTTATGAAACGAGTCAGCTGGATGACCAGAGTCTGAAACC
20 CACAGCCACAAGCAGTCCAGATTATATAAGCGGAAAGCCAAATGATGAGAGCAATGAGCATTCGGATGTGATTGATAGTCAGGAA
CTTTCCAAAGTCAGCCGTGAATTCCACAGCCATGAATTTACAGCCATGAAGATATGCTGGTTGTAGACCCCAAAAGTAAGGAA
GAAGATAAACACCTGAAATTTTCGTATTTCTCATGAATTAGATAGTCATCTTCTGAGGTCAATTAAGAGGAGAAAAATACAAT
TTCTCACTTTGCATTTAGTCAAAAGAAAAATGCTTTATAGCAAAATGAAGAGAACATGAAATGCTTCTTCTCAGTTTATGTG
GTTGAATGTGATCTATTTGAGTCTGGAAATAACTAATGTGTTTGATAATTAGTTTAGTTTGTGGCTTCATGGAACTCCCTGT
25 AAATAAAAGCTTCAGGGTTATGTCTATGTTTCATTCATAGAAGAAATGCAAACTATCACTGTATTTAATATTTGTTATCTC
TCATGAATAGAAATTTATGTAGAAGCAACAAAATACTTTTACCCACTTAAAAAGAGAATATAACATTTTATGTCACTATAATC
TTTTGTTTTTTAAGTTAGTGATATTTTGTGTGATATCTTTTGTGGTGTGAATAAATCTTTTATCTTGAATGTAATAAGAA
TTTGGTGGTGTCAATTGCTTATTTGTTTTCACGGTTGTCCAGCAATTAATAAACATAACCTTTTCTTACTGCTTAAAAAAA
AAAAAAA

30 269

MWTLVSWVALTAGLVAGTRCPDGQFCPVACCLDPGGASYSCRPLLDKWPTTLSRHLGGPCQVDAHCSAGHSCIFTVSGTSSCC
PFPEAVACGDGHHCCPRGFHCSADGRSCFQSRGNSVGAIQCPDSQFECPDFSTCCVMVDGWSGCCPMPQASCCEDRVHCCPHG
AFCDLVHTRCITPTGTHPLAKKLPAQRTNRAVALSSVMCPDARSRCPDGSTCCELPBGKYGCCPMPNATCCSDHLHCCPQDTV
CDLIQSKCLSKENATDILLTKLPAHTVGDVKCDMEVSCPDGYTCRLQSGAWGCCPFTQAVCCEDHIHCCPAGFTCDTQKGTCE
35 QGPHQVPWMEKAPAHLSLPDQALKRDVPCDNVSSCPSSDTCCQLTSGEWGCCPIPEAVCCSDHQHCCPQGYTCVAEGQCQRGS
EIVAGLEKMPARRASLSHPRDIGCDQHTSCPVGQTCPSLGGSWACCQLPHAVCCEDRQHCCPAGYTCNVKARSCEKEVVSQAP
ATFLARSPHVGVKDVCEGEGHFCHDNQTCRRDNRQGWACCPIRQGVCCADRRHCCPAGFRCAARGTKCLRREAPRWDAPLRDPA
LRQLL

270

GTAGTCTGAGCGCTACCCGGTTGCTGCTGCCCAAGGACCGCGGAGTCGGACGCAGGCAGACCATGTGGACCCTGGTGAGCTGGG
TGGCCTTAACAGCAGGGCTGGTGGCTGGAACCGCGTGCCAGATGGTCAGTTCTGCCCCTGTGGCCTGCTGCCTGGACCCCGGAG
GAGCCAGCTACAGCTGCTGCCGTCCCTTCTGGACAAATGGCCCCACAACACTGAGCAGGCATCTGGGTGGCCCCCTGCCAGTTG
ATGCCCACTGCTCTGCCGGCCACTCTGCTCATCTTTACCGTCTCAGGAGCTTCCAGTTGCTGCCCTTCCAGAGGCGCGTGCAAT
GCGGGGATGGCCATCACTGCTGCCACGGGCTTCCACTGCAGTGACAGCGGCGATCCTGCTTCCAAAGATCAGGTAACAACT
45 CCGTGGGTGCCATCCAGTGCCCTGATAGTCAGTTGCAATGCCCGGACTTCTCCACGTGCTGTGTTATGGTCGATGGCTCCTGGG
GGTGTGCCCATGCCACAGGCTTCTGCTGTGAAGACAGGTGCACTGCTGTCCGCACGGTGCTTCTGCGACCTGGTTTCA
CCCGCTGCATCACACCCAGGGCACCCACCCCTGGCAAAGAAGCTCCCTGCCAGAGGACTAACAGGGCAGTGGCTTGTCCA
GCTCGGTGATGTGTCCGACGCACGGTCCCGGTGCCCTGATGGTTCTACCTGCTGTGAGCTGCCACCTGGGAAGTATGGCTGCT

ATCTCTGGACTCCCGGGGAACCCACAGTGGAGGTGGATCTCTATACTGCCAAAGGCTCTTTTCGGGGCTGCAGTGGCCAGTGA
GCCTCTACGGGCATCTATGAGGCCCTGGAGCTGAGGGATGGAGACAAACAGCGTTACTTAGGCAAAGGTGTCTGAAGGCAGTG
GACCACATCAACTCCACCATCGCGCCAGCCCTCATCAGCTCAGGTCTCTCTGTGGTGGAGCAAGAGAACTGGACAACCTGATG
5 CTGGAGTTGGATGGGACTGAGAACAAATCCAAGTTTGGGGCCAATGCCATCCTGGGTGTGTCTCTGGCCGTGTGTAAGGCAGGG
GCAGCTGAGCGGGAACTGCCCTGTATCGCCACATTGCTCAGCTGGCCGGGAACTCAGACCTCATCTGCCTGTGCCGCCCTTC
AACGTGATCAATGGTGGCTCTCATGTCTGCAACAAGCTGGCCATGCAGAGATTATGATCCTCCAGTGGGAGCTGAGAGCTTT
CGGGATGCCATGCGACTAGGTGACAGAGGTCTACCATACACTCAAGGGAGTCATCAAGGACAAATACGGCAAGGATGCCACCAAT
GTGGGGGATGAAGGTGGCTTTGCCCCCAATATCTTGGAGAACAGTGAAGCCTTGGAGCTGGTGAAGGAAGCCATGCACAAGGCT
GGCTACACGGAAAAGATCGTTATTTGGCATGGATGTTGCTGCCTCAGAGTTTTATCGTGATGGCAAATATGACTTGGACTTCAAG
10 TCTCCCACTGATCCTTCCCGATACATCACTGGGGACCAGCTGGGGGCACTCTACCAGGACTTTGTTCAGGGACTATCCTGTGGTC
TCCATTGAGGACCCATTTGACACAGGATGATTGGGCTGCCCTGGTCCAAGTTACACGCCAAATGTAGGGATCCAGATTGTGGGTGAT
GACCTGACAGTGACCAACCAAAAACGTAATTGACGGGGCAGTGGTGAAGAAAAGCGCTGCAACTGTCTGCTGCTCAAGGTCAACCG
ATCGGCTCGGTACTGTAAGCCATCCAAGCGTGCAAGCTGCCCCAGGAGATGGCTGGGGGTCTAGTGGTGAATCATCGCTCAGGA
GAGACTGAGGACACATTTCATTGCTGACCTGGTGGTGGGGCTGTGCACAGGCCAGATCAAGACTGGTGGCCCCGTGCCGTTCTGAA
15 CGTCTGGCTAAATACAACCAGCTCATGAGAATTGAGGAAGAGCTGGGGGATGAAGCTCGCTTTGCCGGACATAAATTCGGTAAT
CCCAGTGTGCTGTGATTCTCTGCTTGGCTGGAGAGCTGGAACCTCTGTCTCATCCTCCTGGAACCTTGTGTCTCTGATCTGTG
ATAGTTACACCCCTGAGATCCCCTGAGCCCCAGGGTGCCAGAACTTCCCTGATTGACCTGCTCCGCTGCTCCTTGGCTTACCT
GACCTCTTGTCTGTCTGTCTCGCCCTCCCTTTCTGTGCCCTACTCATTGGGGTTCCGCACTTTCCACTTCTTCCTTTCTCTTTCT
CTCTTCCCTCAGAAACTAGAAATGTGAATGAGGATTAATTATAAAAGGGGTCCGTGGAAGAATGATCAGCATCTGTGATGGGAG
20 CGTCAGGGTTGGTGTGCTGAGGTGTTAGAGAGGGACATGTGTCACTTGTGCTTTGTCTCTGTGCCACGTGTCTTCCACTTTGC
ATATGAGCCGTGAACGTGTGCATAGTGTGGGATGGAGGGGAGTGTGGGCATGTGATCACGCCTGGCTAATAAGGCTTTAGTGT
ATTTATTTTATTTATTTATTTATTTGTTTTTCATTCATCCCATTAATCATTTCCCCATAACTCAATGGCCTAAAACCTGGCCTGA
CTTGGGGGAAGCGATGTGTCTGTAATTCATGTGGCTGTAGATCCCAAGATGACAGTGGGGTGGGAGGTCTTGTGTAAGATGGGAAGGG
TCTAGAAAAGGCCCTTGACATCACTTTCCTTTGTGTGTAATCACTGTAAGCTGCGTGTGGTCCAGAGCGGAGGCTGTGTGCCCTGGG
25 GGAGTTTTCCTCTATACATCTCTCCCCAACCTAGGTTCCCTGTTTCTTCTCCAGCTGCACAGAGACCTCTCACTCCCAT
GCCACGTTCCACAGTTGCCACCACCTCTGTGGCATTGAAATGAGCACCCTCCATTAAAGTCTGAATCAGTGC

275

30 MPLNRTLSSSLPGLDWEDEFDLENAVLFEEVAWEVANKVGGIYTVLQTKAKVTGDEWGDNYFLVGPYTEQGVRTQVELLEAPT
PALKRTLDSMNSKGCKVYFGRWLIEGGPLVLLDVGASAWALERWKELWDICNIGVPWYDREANDAVLFGFLTTFWLFGEFLAQ
SEEKPHVVAHFHEWLAGVGLCLCRARLPVATIFTTHATLLGRYLCAVDFYNNLENFNVDKEAGERQIYHRYCMERAAAHCA
HVFTTVSQITAEAQHLLKRKPDIVTPNGLNVKKFSAMHEFQNLHAQSKARIQEFVRGHPYGHLDNFNLDKTLTYFFIAGRYEFSN
KGADVLEALARNYLLRVNGSEQTVVAFIMPARTNNFNVETLKGQAVRKQLWDTANTVKEKFGRLKLYESLLVGSPLDMNMKML
DKEDFTMMKRAIFATQRQSFPPVCTHNMMLDDSSDPILTTIRRIGLFNSSADRVKVIHFPEFLSSTSPLLPVDYEEFVRGCHLGV
FPSYYEPWGYTPAECTVMGIPSISTNLSGFGCFMEEHIADPSAYGIYILDRFRSLDDSCSQLTSFLYSFCQQSRRQRIQRNR
35 TERLSDLLDWKYLGRYYMSARHMALSKAPPEHFTYEPNEADAAQGYRPPASVPPSPSLSRHSSPHQSEDEEDPRNGPLEEDG
ERYDEDEEAAKDRRNIRAPEWPRRASCTSSTSGRKRNSVDTATSSSLSTPSEPLSPTSSLGEERN

276

CGCTTCGGGCAGGGGTGCGGTCTTGC AATAGGAAGCCGAGCGTCTTGCAAGCTTCCCGTCGGCACCAGCTACTCGGCCCCGCGAC
CCTACCTGGTGCATTTCCCTAGACACCTCCGGGGTCCCTACCTGGAGATCCCGGAGCCCCCTTCTGCGCCAGCCATGCCTTT
40 AAACCGCACTTTTGCCATGTCTCACTGCCAGGACTGGAGGACTGGGAGGATGAATTGCACTGGAGAACGCAGTGCTCTTCGA
AGTGGCCTGGGAGGTGGCTAACAAGGTGGGTGGCATCTACACGGTGCTGCAGACGAAGCGAAGGTGACAGGGGACGAATGGGG
CGACAAC TACTTCTGGTGGGGCCGTACACGGAGCAGGGCGTCAGGACCAGGTGGAAC TCTGGAGGCCCCACCCCGGCCCT
GAAGAGGACACTGGATTCCATGAACAGCAAGGGGTGCAAGGTGTATTTCGGGCGCTGGCTGATCGAGGGAGGCCCTCTGGTGGT
GCTCTGGACGTGGGTGCCTCAGCTTGGGCCCTGGAGCGCTGGAAGGGAGAGCTCTGGGATATCTGCAACATCGGAGTGCCGTG
45 GTACGACCCGCGAGGCCAACGACGCTGTCTCTTTTGGCTTTCTGACCACTTGGTTCTGGGTGAGTTCTTGGCACAGAGTGAGGA
GAAGCCACATGTGGTTGCTCACTTCCATGAGTGGTTGGCAGGCGTTGGACTCTGCCTGTGTGCTGCCCGGCGACTGCCTGTAGC
AACCATTTCACACCACTGCCACGCTGCTGGGGCGCTACCTGTGTGCGGTGCGCGTGGACTTCTACAACAACCTGGAGAACTT
CAACGTGGACAAGGAAGCAGGGGAGAGGCAGATCTACCAAGCATGCACTGGAAGAGGCGGCGACCCACTCGCGCTGCATCTT
CACTACTGTGTCCAGATCACCGGCATCGAGGCACGACATCTGTCTAAGAGGAAACAGATATTGTGACCCCAATGGGCTGAA
50 TGTGAAGAAGTTTTCTGCCATGCATGAGTTCCAGAACCTCCATGTCTCAGAGCAAGGCTCGAATCCAGGAGTTTGTGCGGGGCCA

TTTTTATGGGCATCTGGACTTCAACTTGGACAAGACCTTATACTTCTTTATCGCCGGCCGCTATGAGTTCTCAACAAGGGTGC
TGACGTCTTTCTGGAGGCATTGGCTCGGCTCAACTATCTGCTCAGAGTGAACGGCAGCGAGCAGACAGTGGTTGCCTTCTTCAT
CATGCCAGCGCGGACCAACAATTTCAACGTGGAACCCCTCAAAGGCCAAGCTGTGCGCAACAGCTTTGGGACACGGCCAACAC
GGTGAAGGAAAAGTTCGGGAGGAAGCTTTATGAATCCTTACTGGTTGGGAGCCTTCCCGACATGAACAAGATGCTGGATAAGGA
5 AGACTTCACTATGATGAAGAGAGCCATCTTTGCAACGCAGCGGAGTCTTTCCCCCTGTGTGACCCACAATATGCTGGATGA
CTCTCAGACCCCATCCTGACCACCATCCGCCGAATCGGCCTCTTCAATAGCAGTGCAGCAGGGTGAAGGTGATTTTCCACCC
GGAGTTCTCTCTCCACAAGCCCCCTGCTCCCTGTGGACTATGAGGAGTTTGTCCGTGGCTGTACCTTGGAGTCTTCCCTC
CTACTATGAGCCTTGGGGCTACACACCGGCTGAGTGCACGGTTATGGGAATCCCAGTATCTCCACCAATCTCTCCGGCTTCGG
CTGCTTCATGGAGGAACACATCGCAGACCCCTCAGCTTACGGTATCTACATTCTTGACCGCGGTTCCGGCAGCTGGATGATTC
10 CTGCTCGCAGCTCACCTCCTTCTCTACAGTTTCTGTGTCAGCAGAGCCGGCGGACGCTATCATCCAGCGGAACCGCAGGAGCG
CCTCTCCGACCTTCTGGACTGGAATACCTAGGCCGGTACTATATGTCTGCGCGCCACATGGCGCTGTCCAAGGCCTTTCCAGA
GCATTTCACTACGAGCCCAACGAGGCGGATGCGGCCAGGGGTACCGTACCCACGGCCAGCTCGGTGCCACCGTCCGCCCTC
GCTGTACGACACTCCAGCCCGCACCAGAGTGAGGACGAGGAGGATCCCGGAACGGGCCGCTGGAGGAAGACGGCGAGCGCTA
CGATGAGGACGAGGAGGCCGCCAAGGACCGGCCAACATCCCTGCACCCAGAGTGGCCGCGCCGAGCGTCTGCACCTCCTCCAC
15 CAGCGGCCGAAGCGCAACTCTGTGGACACGGCCACCTCAGCTCACTCAGCACCCCGAGCGAGCCCCCAGCCCCACCGACTC
CCTGGGCGAGGAGCGTAACCTAAGTCCGCCCCACACACTCCCGCCTGTCTGCTCTCTGCTCCAGAGAGGAGTGCAGAGGG
GTGCTGCTCTAAACCCCGATCCAGATCTGCACGGGTGCGGCCCGCAGTGCACCCACCCAGTCCGCCAAACACTCCACCC
CTCCAGCTCCAGTTTCCAAGTCTCTGCACCTTGAATCCACAAGCCGTGCTTTCTCTGGCTCCAGAATATGCATAATCAGCG
CCCTGGAGTCCCTTGGGCTTGGACCGCTTCCAGAGGCCAGGAATCTGCCATTACTCTGCGGTGGTGCAGAGGTTTATAGGAA
20 ACCTGGCATGGTGTCTTTCAGGCTTGGGGCTTTTAGAGCCCCCGTGTGGCTTACAAATTTACAGCATAACAGAGCAGGCCACGC
TCAGGCCCGCATGCGGGCCACCAAGTTCTGGAACACCGTGGTGTCCCTGCGAATGGGGCGATCAAGTCCAGAGCCGGGGCAC
TTTCAGAGTTTGAAGGTAACCTGAGAGCAGATGGTCTCCATTTCAACTCCAGAAGTGGGGCTCTGGGAGGGATGTTCTAGCCCT
CCCTGGCTGTGAGAGCCAGGCTCTGCTGGAGGATCCCTCCATCCGGCTCCTGTCTATCCCTTACACTTTGGCCAAGCAAGAGGT
GGTAGAACCACTTGGCTGCTCATCTCTTCTGGAGGACACAGTCTCAGTCCAGATGCCTTCTGTCTTCTTGTGCTCTTCTGG
25 ACCAGATCTACTCTTCTTTCTAAATCTGAGATCTCCCTCCAGGGAATCCGCCCTGCAGAGGACAGAGCTGGCTGTCTTCCCC
ACCCCTAACCTGGCTTATTTCCCACTGCTCTGCCCACTGTGAAACCACTAGGTTCTAGGTCTTGGCTTCTAGATCTGGAACCTT
ACCACGTTACTGCATATGATCCCTTTCCCATGATCCAGAACTGAGGTCACTGGGTCTTAGAACCCCCACATTTACCTCGAGGC
TCTTCCATCCCCAACTGTGCCCTGCCTTTCAGCTTTGGTGAAGGGAGGGCCCCATGTTGTGCTGTGCTGTGCTGTGCTGCACCGCT
TGGTTTGCAGTTGAGAGGGGAGGGCAGGAGGGGTGTGATTGGAGTGTGTCCGGAGATGAGATGAAAAAATACATCTATATTTA
30 AGA
277
MESKGASSCRLLFCLLISATVFRPGLGWYTVNSAYGDTIIIPCRLDVPOQLMFGKWYKPKDGPVFIAPRSSTKKSQVYDDVP
EYKDRNLNSENYSISNARISDEKRFVCMVLTEDNVFEAPTIVKVFQPSKPEIVSKALFLETEQLKKLGDCISEDSPDGN
TWYRNGKVLHPLEGAVVIIFFKEMDPVTQLYTMTSTLEYKTTKADIQMPFTCSVTYYGPSGQKTIHSEQAVFDIYYPTEQVTIQ
35 VLPPKNAIKEGDNITLKLNGNPPPEEFLFYLPQPEGIRSSNTYTLMDVRRNATGDYKSLIDKKSMAISTAITVHYLDLSL
NPSGEVTRQIGDALPVSCITISARNATVVMKDNIRLRSSPSFSSLHYQDAGNYVCETALQEVGLKKRESLTLIVEGKPKQIKM
TKKTDPGSLSKTIIICHVEGFPKPAIQWTITGSGSVINQTEESPYINGRYYSKIIISPEENVTLCTAENQLERTVNSLNVSAIS
IPEHDEADEISDENREKVNDQAKLIVGIVVGLLLAALVAGVVYLYMKKSKTASKHVNKDLGNMBENKLEENHNKTEA
278
40 CGGGACGACGCCCCCTCTGCGGCGTGGACTCCGTCAGTGGCCACCAAGAAGGAGGAGGAATATGGAATCCAAGGGGGCCAGT
TCCTGCCGTCTGCTCTTCTGCCTCTTGATCTCCGCCACCGCTTTCAGGCCAGGCCTTGGATGGTATACGTAAATTCAGCATAT
GGAGATACCATTTATCATACCTTGGCGACTTGACGTACCTCAGAATCTCATGTTTGGCAAATGGAATATGAAAAGCCCGATGGC
TCCCCAGTATTTATTTGCTTTCAGATCTCTACAAAGAAAAGTGTGCAGTACGACGATGTACCAGAATACAAAGACAGATTGAAC
CTCTCAGAAAACFACACTTTGTCTATCAGTAATGCAAGGATCAGTGATGAAAAGAGATTGTGTGCATGCTAGTAACCTGAGGAC
45 AACGTGTTTGGAGCACCTACAATAGTCAAGGTGTTCAAGCAACCATCTAAACCTGAAATGTGAAGCAAAAGCATTGTTCTCGAA
ACAGAGCAGCTAAAAAAGTTGGGTGACTGCATTTTCAAGACAGATTATCCAGATGGCAATATCAGATGGTACAGGAATGAAAAA
GTGTACATCCCTTGAAGGAGCGGTGGTCATAATTTTTTAAAAAGGAAATGGACCCAGTCACTCAGCTCTATACCATGACTTCC
ACCTTGGAGTACAAGACAACCAAGGCTGACATACAAATGCCATTACCTGCTCGGTGACATATTATGGACCATCTGGCCAGAAA
ACAAATTCATTCTGAACAGGCAGTATTTGATATTTACTATCTACAGAGCAGGTGACAATACAAGTGTGCCACCAAAAAATGCC
50 ATCAAAGAAGGGGATAACATCACTCTTAAATGCTTAGGGAATGGCAACCCCTCCCCAGAGGAATTTTGTGTTTACTTACCAGGA

MERVKMINVQRLLLEAAEFLERRERERECEHGYASSFPSPSPRIQHSKPPRLSRAQKHSSGTSNTSTANRSTHNELEKNRRAHLR
 LCLERLKVLIPLGPDCTRHTTLGLLNKAKAHIKKLEEAERKSQHQLLENLEREQRFLKWRLEQLQGPQEMERIRMDSIGSTISSD
 RSDSREEIEVDVESTEFSHGEVDNISTTSISDIDDHSSLPISIGSDEGYSSASVKLSFSS

[illegible]

CAGCACTTGTCTCATTTTAAATGTAAAGATTGCTTCCATTTTCTTACAGGCAGTCTCTCTCTCCTCACAGTCCCAGTGTGCAG
GTGCTATTGTTACTCTTACGAATATTTTTCAGTAATGTTATTTTCTTCTAAGTGAAATTTCTAGCCTGCAGTTTGATGTCATGTG
TTCCCTTTGTCTTTCAAACCTCCAAGGTTCCCTGTGGCCCTCTCCCTTACCTTGGGAAGGCCCTCTTGAGACCTTACCCCTGGC
TGTGTGGACTTTGTATACTTTAAATAATTTAACTACCCTTAATTACTTAAAAA

5 282

ATCGCAACCTGGGCCTCAGCATGCTGCTGCAGTGATGGCGCCAACCCGTGGCAGTCCCAGAGCTGGAGGCAGGAGGATGGATCC
TCATCTCCATGGGAAGTGTACAGCTGTGGCTGCCAGGAAGCGTGGCAGGCGCTGGCCTTGGGTCCATCTACATAGTTGCGTG
TTTCAACAATGTCCATTTATCCTTACCCCGAGGCGTGTGTTGGGGCTGCAAAACCTCCCGGTAGAGGCTGGACCTGAGGAC
CCTTCCCACCTGTGCCCCCTTCTGAAAGTCTTAGCCACAGCCATCCTCCATGAGTCCCGGCAGCTCTGGGTGATGCCCCCT
10 CCTGGTCACCCATCTGCCCTCACCTCGTCATCCAGGACCCAGACCTGCACCTTCCATGTGGGCCACAGATCCTTGGCAG
TGACCTGAGGTGCACCATTGAGTGTGCGATTGGGGTTAGCATCCAGAAAGAAGAATGCGCATGACGCTCTGTGAAGGCTGGAA
CTCAGGCTCTCAGGGAGAGAAAGGAAGACTGGATTGCACCTTGATGCCCTCTGAGGAGCGGCCCTCTTGTGAGGTGGCGTG
GGCCCGGCCAGCCTTATCCAAGTCGCTCTGTCCACCTCCCTTCTGCCCCCACCCTCTGTGCTCCAGGAGCCCT
CCCTGTGCTCCACCTGCCCTCCGAGAAGGAAGCCTCTTCTCTGTTTCCCTGGGTGAGGGGGTGGCAGGTGGCTAACCCCAT
15 TAGCATCTCCAGGCCCTGCCATGGTGTCTCATCTTGTGTTATCTCTAGCTCTTCCCTCTCCCATTTCTTTAGTAGTTGAA
TTTTGCAAAGCTTGTAGCAGTAGCTCAGTTGCCGTCAGCATCCTTGTGTGTAGATAAATAGTCGACAGAACTCAGCACTGGG
GACAGGATTGCAAAGTCGGGGACATAGATGCAGACAGTTGTTGAGATTGGGGATAGCCGGGCTGTGAGCGGTGCCCATTTCC
AGATGAAGCCTTTAGCCCTTCTGAGTCCCCGGCCCTGGTGCGATGTCTGTGAGTTTGACCTGCCAGCGTGTGGGCTGGCTC
AATGCTGAATAAAGTGGGTTGTGTCAAAAAAAAAAAAAAAAAAAAAA

20 283

MDQKSLWAGVVVLLLLQGG SAYKLVCYFTNWSQDRQEPGKFTPENIDPFLCSHLIYSFASIENNKVIIKDKSEVMLYQTINSLK
TKNPKLKILLSIGGYLFSGKGFHPMVDSSSTRLEFINSIILFLRNHNPDLVDVSWIYPDQKENTHFTVLIHELAEAFQKDFTKS
TKERLLLTAGVSAGRQ MIDNSYQVEKLA KDLDFINLSPDFHGSWEKPLITGHNSPLSKGWQDRGPSSYINVEYAVGVYIHKGM
PSEKVVMIPTGYHSFTLASAETTVGAPASGPGAAGPITBSSGFLAYYEICQFLKGAKITRLQDQVPYAVKGNQWVGYYDDVKS
25 METKVQFLKNLNLGGAMIWSIDMDDFTGKSCNQGPYPLVQAVKRSLSL

284

AGAAGAAGCTGGCCAAGGATATGGGAGCAACCACCATGGACCAGAAGTCTCTCTGGGCAGGTGTAGTGGTCTTGTCTGCTTCTCC
AGGGAGGATCTGCCACAACTGGTTTGCTACTTTACCAACTGGTCCCAGGACCGGCAGGAACCAGGAAATTCACCCCTGAGA
ATATTGACCCCTTCTATGCTCTCATCTCATCTATTTCATTCGCCAGCATCGAAAACAAGGTTATCATCAAGGACAAGAGTG
30 AAGTGATGCTCTACCAGACCATCAACAGTCTCAAAACCAAGAATCCAACTGAAAATTTCTTGTCTCATTTGGAGGGTACCTGT
TTGGTTCCAAAGGGTTCCACCTATGGTGGATTCTTCTACATCAGCTTGAATTCATTAACCTCATTAATCTGTTTCTGAGGA
ACCATAACTTTGATGGACTGGATGTAAGCTGGATCTACCCAGATCAGAAAGAAACACTCATTTCACTGTGCTGATTCTATGAGT
TAGCAGAAGCCTTTGAGAAGACTTCACAAAATCCACCAAGGAAAGGCTTCTCTGACTGCGGGCGTATCTGCAGGGAGGCAAA
TGATTGATAACAGCTATCAAGTTGAGAACTGGCAAAAGATCTGGATTTTCATCAACCTCTGTCTTTGACTTCCATGGGTCTT
35 GGGAAAAGCCCTTATCACTGGCCACAACAGCCCTCTGAGCAAGGGGTGGCAGGACAGAGGGCCAAGCTCTACTACAATGTGG
AATATGCTGTGGGGTACTGGATACATAAGGAATGCCATCAGAGAAGGTGGTCAATGGGCATCCCCACATATGGGCACTCCTTCA
CACTGGCCTCTGCAGAAACACCGTGGGGGCCCCTGCTCTGGCCCTGGAGCTGCTGGACCCATCACAGAGTCTTCAGGCTTCC
TGGCCTATTATGAGATCTGCCAGTTCTTGAAGGAGCCAAGATCAGCGCCTCCAGGATCAGCAGGTTCCCTACGCAGTCAAGG
GGAACCAGTGGGTGGGCTATGATGATGTGAAGAGTATGGAGACCAAGGTTAGTTCTTAAAGAATTTAAACCTGGGAGGAGCCA
40 TGATCTGGTCTATTGACATGGATGACTTCACTGGCAAATCTGCAACCAGGGCCCTTACCCTCTGTCCAAGCAGTCAAGAGAA
GCCTTGGCTCCTTGTGAAGGATTAACTTACAGAGAAGCAGGCAAGATGACCTTGTCTGCTGGGGCTGCTCTCTCCAGGAATT
CTCATGTGGGATCCCCCTTGCCAGGCTGGCCTTTGGATCTCTCTTCCAAGCCTTTCTGACTTCTCTTAGATCATAGATTGGA
CCTGGTTTTGTTTTCTGTCAGCTGTTGACTTGTGGCCCTGAAGTACAATAAAAAAATTCATTTTGTCTCAGTA

285

45 MSLSENSVFAYESSVHSTNVLLSLNDQRKKDVLCDVTIFVEGQRFRAHRSVLAACSSYFHSRIVGQADGELNITLPEEVTVKGF
EPLIQFAYTAKLILSKENVDEVCKVEFLSVHNIEESCFQFLKFKPLDSTADQEQCPKKCFSSHCKQKDLKLSLLDQRLDLET
EVEEFLKNVQTPQCKLRRYQGNASPPQDSASQTYESMCLEKDAALALPSLCPKYRKFKQAFGTDRTVTESSVKDIHAS
VQPNERSENECLGGVPECRDLQVMLKCDESKLAMEPEETKKDPASQCPTEKSEVTPFPHNSSIDPHGLYLSLLHTYDQYGD LN

FAGMQNTTVLTEKPLSGTDVQEKTFGESQDLPLKSDLGTREDSSVASSDRSSVEREVAEHLAKGFWSDICSTDTPCQMQLSPAV
AKDGSEQISQKRSECPWLGI RISESPEPGQRTFTTLSSVNCPISTLSTEGCSSNLEIGNDDYVSEPQQEPCPYACVISLGDDSD
ETDTEGDSESCSAREQECEVKLPFNAQRIISLSRNDFFQSLKMHKLTPEQLDCIHDRRRSKNRIAAQRCRKRKLDCIQNLESE
IEKLQSEKESLLKERDHLSTLGETKQNLTLGLCQKVCKEAAALSQEQIILAKYSAADCPLSFLISEKDKSTPDGELALPSIFSL
5 SDRPPAVLPPCARGNSEPGYARGQESQQMSTATSEQAGPAEQCRQSGGISDFCQOMTDKCTTDE

286

TCGCCCCCGCGGGCGCTCTCGCTTCAGTCAGTCGGGCCGCGCGCGCCTCAGCTCTGGTTGATGATAATTAGAAGCATGCTTT
CCACTGAACCTCCCGACACACATTTGTTATGACAGAAATGCTCTGAGTGAGAACTCGGTTTTTGCCATATGAATCTTCTGTGCATAG
CACCATGTTTTACTCAGCCTTAATGACCAGCGGAAGAAAGATGTGCTGTGCGATGTCACCATCTTTGTGGAGGGACAGCGGTT
10 CCGCGCTCACCGGTCGCTGCTGGCGGCATGCAGCAGTTACTTCCACTCAAGAAATCGTAGGCCAGGCTGATGGAGAGCTGAACAT
TACTCTTCCAGAAGAGGTGACAGTTAAAGGATTTGAACCTTTAATTCAGTTTGCTTACACTGCTAAACTGATTTTAAAGTAAAGA
GAATGTGGATGAAGTGTGCAATGTGTGGAGTTTAAAGTGTACATAATATTGAGGAATCCTGCTTTTCAAGTTTCTGAAATTTAA
GTTTTTGGACTCCACTGCAGACCAGCAAGAAATGCCCAAGAAAAAATGCTTTTCATCACACTGTCAGAAAAACAGACCTTAAACT
TTCACTTTTGGACCAGAGGATCTAGAACTGATGAAGTGGAGGAATTTCTGGAAAAATAAAATGTTTCAAGTCTCTCAGTGTA
15 ACTCCGCAGGTATCAAGGAAATGCAAAAGCCTCACCTCTCTACAAGACAGTGCCAGTCAGACATATGAGTCCATGTGCTTAGA
GAAGGATGCTGCTCTGGCCTTGCCTTCTTTATGCCCAAATACAGAAAAATCCAAAAAGCATTTGGAAGTACAGAGTCCGTAC
TGGGGAATCTAGTGTCAAAGACATTCATGCTTCTGTTTACGCCAAATGAAAGGTCTGAAATGAATGCCCTGGGAGGAGTCCCGGA
GTCTAGAGATTTGACAGGTGATGTTAAATGTGACGAAAGTAAATTAGCAATGGAACTGAAGAAACGAAGAAAGATCCTGCTTC
TCAGTGCCCAACTGAAAAATCAGAAAGTACTCCTTTCCCCACAAATTTCCATAGACCTCATGGACTTATTTCTTTGTCTCT
20 TTTACACACATATGACCAATATGGTGACTTGAATTTTGTGTTATGCAAAACACAAACAGTGTAAACAGAAAAAGCCTTTGTGAGG
TACAGACGTCCAAGAAAAACATTTGGTGAAAGTCAGGATTTACCTTTGAAATCCGACTTGGGCACCAGGGAAGATAGTAGTGT
TGCATCTAGTGATAGGAGTAGTGTGGAGCGAGAAGTGGCAGAACACCTAGCAAAAGGCTTCTGGAGTGACATTTGCAGCACGGA
CACTCCTTGCCAAATGCAGTTATCACCTGCTGTGGCCAAAGATGGCTCAGAACAGATCTCAGAAACGGCTGTGAGTGCCGTG
GTTAGGTATCAGGATTAGTGAGAGCCAGAACAGGCTCAAAGGACTTTCAACATTAAGTCTGTCAACTGCCCTTTTATAAG
25 TACTCTGAGTACTGAAGGCTGTTCAAGCAATTTGAAATTTGAAACGATGATTATGTTTCAAGACCCAGCAAGAACCTTGCC
ATATGCTTGTGTCATTAGCTTGGGAGACGACTCTGAGACGGACACCCGAAGGAGACAGTGAATCCTGTTTCAAGAGAACAAAGA
ATGTGAGGTAAAACTGCCATTTCAATGCAACACGGATAATTTTCACTGTCTCGAAATGATTTTCAAGTCTCTTGTGAAAAATGCAAA
GCTTACTCCAGAACAGCTGGATTGTATCCATGATATTCGAAGAAGAAGTAAAAACAGAAATTTGCTGCACAGCGCTGTGCAAGAG
AAAATTTGACTGTATACAGAACTTTGAATCAGAAATTTGAGAAGCTGCAAAAGTGAAGGAGAGCTTGTGAAAGGAAAGAGATCA
30 CATTTTGTCAACTCTGGGCGAGACAAAGCAGAACCTTAAGTACTTTGCGCAGAAAGTTGTAAAGAGCAGCTCTGAGTCAAGA
ACAAATACAGATACTCGCCAAGTACTCAGCTGCAGATTGCCACTTTTCAATTTTAAATTTCTGAAAAAGATAAAAGTACTCTGA
TGGTGAAGTGGCGTTACCATCAATTTTCAAGTTTATCTGACCGGCTCCAGCAGTGTGCTCCCTGTGCGCAGAGGAAACAGTGA
GCCTGGCTACGCGGAGGGCAGGAGTCCAGCAGATGTCCAGACCCACCTCTGAGCAAGCTGGGCTGCGGAACAGTGTGCTCA
GAGTGGTGGATCTCAGATTTCTGTGTCAGCAGATGACTGATAAATGTACTACTGATGAGTAAACTTGCATTACCTTCTTCAAC
35 CATCTAATTTTCTCCTGAAAGTTTGGCAGCGCTTTGAAAGCCTAATATGACCATCTGTTGCTCAACAATACTGTTTCTTCTCT
TAGTAGTTTACCACAAGGGAATTTCTTTAAGTCAACCATGATTTCTCTCTGATTTCTACAAGAGACAAAGAAATGATTTTGCC
TCCTGGATATCAGAAAAATCCATGTGAAATGTAGTAAACCTTTAAACTCATGTTTAAAGAAATAAATACTCTAGTAAATACT
CTTCTGCTATTGAGAATAAGTAGGAGAAATGAAACTGCAGCATATCAGACAGCAATTTAAACAGCTTGAACATCTACAGATAG
TTCTACTAAAAGAAAGTGGCCTGCAGAAATTAATAATTTGACTTTTTTCTAATATTTAGTTTGAAGAAAAATTTCTTCCAA
40 GCAATGCTAATAGAGTTCTATTCTTTAGAAGCAGGCTGTGAGTACTGGGAATATTTTGTAGAGCTGCATTGTGAAAAAAGAT
GGTCTTACCTGAATCTTAGGGCTTTGTTCTTCGGCTCTTAAATCAGGCTTTAAGCTACATTGGGAAGATTTAGTAAATAGGCA
AGTGGTTGGCCTAAGACGGGGCTGCTTCTCTCTCAGTATGGACTCTAGAAAGTCTGGCTACATGAATAGATTTAAGTGTCA
CTTTCCCTCCCTGCCCCCGCTTCAGTCTCTACCATATCTGGTCCCATCATGGACTTCTTATTCTCTGGCATTTTGTCCCTTT
GGAAGAAGAAATAGGACTCAGAAATACAGTGGCATGAGTGTACTGAGCAGCATTTATCTCAGGCTCCCTAGAAATCTGGAGAGC
45 TTACCAACATGTAAAGCTGTTTCAATTTTCCACCGTGGGTACCAATGCCAGAAAACAGACATCAGGGGAAAGAAATGTTGCTT
ACTTTTACCAGGAGTGCAGTTCAATTTTTTACCTGTTTTTGAAGTCGTATTATTCAGTTGTAAAAATGATTGTAACAGATA
AAAAATGTATCTGCAGCAACTCTGCAGGTTTGTGAAATAGGATGAAACTCAATCTTTTTCTATTGTGGGTTTGCAATTGAAAG
CAGGTTGAATCCTTGTCTCTTCTCCAAATTTGGTGTGGTATAAAGACACACAAATCATTTTAACTTGGACATTTAAAGATCAG
TCTTAGTGTGTTGTTGAGTCTCTTACAAATAGATAACTGAGCACCTATCGCATAACATTTTGGGTTGGCTTTTAGCCATGCTG
50 GGGTTAGATGTGTTGAGAGTCAATGAAAGCTATGGATCTTCTCAGCAATTAATAAATGCAATATATTCACATTCACAGAAA
CATTTGCAGAACCCAGTTTAAATGGTACAGAGGAGTAGTTTATAGTGTGATTTACCAAAATCAGAGGCTGAAAGAGACACT

TCTATAGACTGCATCCTGAGCCTAGTGCAGGGCTTGTCTAGCTAATGTGGGCAGCCACCACCCACTGTGTATGAACAAGTCTGA
AGCAAGTTGGCCTTGCCCTTGAGAGTATATGGGGACCAGTCTTCATGTCTTGGAGTAATTTGTCAAATGTTACCCCTTTTGTATC
AGGGTGTAGGGGGAGGATATGTCTAGTATATTTTTCAGTGGTTTGTATGTTCTCTGTCTACTGACTTATTTGTAAGAGAAAATT
AGTTGGACTTGTATTTTCTAGTAGCTTTTATAAGTACACTCAAGAATTTGTTCAGGGAGAATAATTTGTATAGTGCATCCCAT
5 ACTGCAAAAGAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTATGTATACATATATATCTCTCCATATAGGTATTTT
TTTGATACTTGTAATTTTAAATTTTCAGCTTCACGATATAAAATAATAAAGAACTTCTGGTTTACAAAATGTAAATCTTAAGC
CAATGGAACCTTGATTTCCCTACCTCAGTGTACACTCAACTATTGGTTGTATCAGTTTGTGTATGTGCAAAATGTCAAATAATCT
TTTGCTTTAATTGCTACTGTACTTGTCTTGAAGATTACCTACTATTTTATGATAAAATGTAGTTGTCTCCAGAGCTTAAATAT
AATTTGTAAAGCACTTGGTTTAAATTTCTCTCTACCTATAAACAGTTTAGCATTAAGGGTTCTATTAATGACACAGAATTATT
10 GGCCAAGTGTAAATTTCTTAAATTTAGCATTACTTTAAATAGCCAGCATGTAATACAAGTAACACTACCTCATATCTACAT
TTGTCATAGCATTACACATTTATGCCATTTTAAACATTAACCTTCTAAAGAACTTTTCTTAAGAAAATGTTTCAAGGCAATATTT
TTTTTGAGGCTGCCGAAGACAAATGACAGGATTATGAGTATACAGTGTATGCCCTTTCTTCATGCAGAATTTTGAATGTTTT
CAGTTTGTATATTGCATATTCACATGATCATGTCTCACTATTTTATGAACCTGGCCTTCTCAATGTTTGTATGATTTTTTAAAGC
15 TGTATGTTGAATTCAGTAAAATAACATTACCTTATTTTTTTCTTATTCAAATTCGGAATATAGCAAAATAATCGTTAAAT
TGTCATATTCAAACAAATGTGGATACAGCTTGGTTCTCCATCTGTAATTTTTTTTAAACAGTTTGTATAGCTTACTGCTTAA
CTAATTTTAAATAAGGAAATAAGTATGTTAGATGCAGTAGACGATACAGGTTGCATGTGGACACTCAGTCACATTAACAACCTG
GGAAAAAATGGCAATGTTACGGTGAATTCACAGTGAACTTTTTTCAGTTATAAAACATCTATTTTGAATCTGTAAATATTTT
AAATGTTTTATTAAGGCATGTAATAAATCTTTTGAACCTGTTGGGTAGAATGAAAATAAGCCATAATGGTAAAGATG
20 GCATACTGATTATAAAAGAAGCAGAAAAACATTGATTTTTTATATCTTTTATAATATAATTTTCTAACAATGCAATAAAACCA
CTAAACTTTTGTGTC

287

MVKIVTVKTQAYQDQKPGTSGLRKRVKVFQSSANYAENFIQSIISTVEPAQRQEATLVVGGDGRFYMKEAIQLIARIAAANGIG
RLVIGQNGILSTPAVSCIIRKIKAIIGGIILTASHNPGGPNDFGIKPNISNGGPAPEAITDKIFQISKITIEYAVCPDLKVDLG
25 VLGKQQFDLENKFKPFTVEIVDSVEYATMLRSIFDFSALKELLSGNRLKICIDAMHGVVGPYVKILCEELGAPANSAVNCV
PLEDFGGHHPDNLTYAADLVETMKSGEHDFAAFDGDGDRNMLILKGHFFVNPSPDSVAVIAANIFSIPIYFQQTGVRGFAFSMP
TSGALDRVASATKIALYETPTGWKFFGNLMDASKLSLCEGESFGTGSDDHIREKDLWAVLWLSILATRKQSVEDILKDHQKH
GRNFFTRYDYEEVEAEGANKMMKDLEALMFDRSFVKGQFSANDKVYTVKEADNFEYSDPVDGSI SRNQGLRLIFTDGSRIVFL
SGTGSAGATIRLYIDSYEKDVAKINQDPQVMLAPLISIALKVSQQLBERTGRTAPTIVT

30 288

GGGCCGCGCCCTCCGCCAGCCAAGTCCGCGCTCTGACCCCCGGCAGCAAGTCGCCACCATGGTGAAGATCGTGACAGTTA
AGACCCAGCGCTACCAGGACCAGAAGCCGGGCACGAGCGGGTGCAGGAGCGGGTGAAGGTGTTCCAGAGCAGCGCCAACTACG
CGGAGAACTTCATCCAGAGTATCATCTCCACCGTGGAGCCGGCGCAGCGGCAGGAGCCACGCTGGTGGTGGCGGGGACGGCC
GGTCTACATGAAGGAGGCCATCCAGCTCATCGCTCGCATCGCTGCCGCCAACGGGATCGGTGCTTGGTTATCGGACAGAATG
35 GAATCCTCTCCACCCCTGCTGTATCCTGCATCATTAGAAAAATCAAAGCCATTGGTGGGATCATTCTGACAGCCAGTCACAACC
CAGGGGGCCCCAATGGAGATTTTGAATCAAAATTCATATTTCTAATGGAGGTCTGCTCCAGAAGCAATAACTGATAAAATTT
TCCAAATCAGCAAGACAATTGAAGAATATGCAGTTTGCCTGACCTGAAAGTAGACCTTGGTGTCTGGGAAAGCAGCAGTTTG
ACTTGGAAAATAAGTTCAAACCTTCACAGTGGAAATGTGGATTCCGTTAGAAGCTTATGCTACAATGCTGAGAAGCATCTTTG
ATTTACAGTGCAGTGAAGAATACTTTCTGGGCCAAACCGACTGAAGATCTGTATTGATGCTATGCATGGAGTTGTGGGACCGT
40 ATGTAAGAAGATCCTCTGTGAAGAATCGGTGCCCCGCGAACTCGGCAGTTAACTGCGTTCTCTGGAGGACTTTGGAGGCC
ACCACCTGACCCCAACCTCACCTATGCAGCTGACCTGGTGGAGACCATGAAGTCAGGAGAGCATGATTTGGGGCTGCCTTTG
ATGGAGATGGGGATCGAAACATGATTCTGGGCAAGCATGGGTCTTTGTGAACCTTCAGACTCTGTGGCTGTCTATGCTGCCA
ACATCTTCAGCATTCGGTATTTCCAGCAGACTGGGGTCCGCGCTTTGCACGAGCATGCCACGAGTGGTGTCTGGACCGGG
TGCTAGTGCTACAAAGATTGCTTTGTATGAGACCCCACTGGCTGGAAGTTTTTGGGAATTTGATGGACGCGAGCAAACTGT
45 CCGTTTGTGGGAGGAGAGCTTCGGGACCGGTCTGACACATCCGTGAGAAAGATGGACTGTGGGCTGTCTTGGCTGGCTCT
CCATCCTAGCCACCGCAAGCAGAGTGTGGAGGACATTCTCAAAGATCATTGGCAAAAGCATGGCCGGAATTTCTTACCAGGT
ATGATTACGAGGAGGTGAAGCTGAGGGCGCAAAATAATGATGAAGGACTTGGAGGCCCTGATGTTTGTATGCTCTTTGTGG
GGAAGCAGTTCTCAGCAAAATGACAAAGTTTACACTGTGGAGAAGGCCGATAACTTTGAATACAGCGACCCAGTGGATGAAGCA
TTTCAAGAAATCAGGGCTTGGCGCTCATTTTACAGATGGTTCTCGAATCGTCTTCCGACTGAGCGGCACTGGGAGTGCCGGG
50 CCACCATTCGGCTGTACATCGATAGCTATGAGAAGGACGTGCGCAAGATTAACCAGGACCCCAAGGTGATGTTGGCCCCCTTA

TTTCCATTGCTCTGAAAGTGTCCTCCAGCTGCAGGAGAGGACGGGACGCACTGCACCCACTGTTCATCACCTAAGAAGACAGGCCTG
ATGTGGTACGTCCCTCCACCCCGGACCCATCCAAGTCATCTGATTGAAGAGCATGACAGAAACAAAATGTATTACCAAGCAT
TTTAGGATTTGACTTTTTCACCTAACCACTTACGAGCAGTGCATTTACAAGGCACCTGCCAAACAAGATGCCCTTGGGAGCTGTG
AGGGAAAAGAGGACCTGCGGGCTTAGATCAATCTCAATTCCTTTTCATGCCCTCCTGCATTGCTGCTGCGTGGGTATTTGTCTCC
5 TTAGCCATCAGGTACAGTTTACACTACAATGTAAGCTATAGGTGGAGCATCAGCAGTGAGTGAGGCCATTCTTCATCCTTAGGA
TGTGGCAATGAAATGATGGTGCAAGTTCCTTTCTCTTTTGTGAATCTTTCCCCCATTTCTGTTTACATGTAACCAACAAAA
TGCAATTTCTAGTGCCCTTCTGTCCAATCAGTTCTTTCTCTGAGTGAGACGTACTTGGCTACAGATTTCTGCCCTGTTTTCGCA
CATTTGCTCCATTACACAGATATTTTGGGATAATAAAGGAAAAATAAGCTACA

289

10 MGSTVPRASVLLLLLLLLRRAEQPCGAELTFELPDNAKQCFYEDIAQGTKSTLEFQVITGGHYDVCRLDPDGKVLKEMKKQ
YDSFTFTASKNGTYKFCFSNEFTFTHKTVYDFQVGETHLCLVDRVSALTQMESACVSIHEALKSVIDYQTHFLREAQGRS
RAEDLNTVRVAYWSVGEALILLVVSIGQVFLKSFSDKRTTTRVGS.

290

GCCTGAGCCGCGCGCTTCCGGAAGCGCAGACCCGCTGGTGCCACGTTTATCCCTTACATCCTCTAGGACCCGGTTCGGTA
15 GTCGTCGCCCCAGCCCGCGGGGCAAGCGCCAGGTCCGCGGCCCTCGAGACGGGACCGAGAGCATCATGGGAGCACTGTCC
CGCGCTCCGCTCCGTGCTGCTTCTGCTGCTGCTCTGCGCGGGCCGAGCAGCCCTGCGGGGCGAGCTCACCTTCGAGCTGC
CGGACAACGCCAAGCAGTGCTTCTACGAGGACATCGCTCAGGACCAAGTCGACCTGGAGTTCCAGGTGATTACTGGTGGTC
ACTATGATGTAGATTGTCGATTAGAAGATCCTGATGGTAAAGTGTTATACAAAGAGATGAAGAAACAGTATGATAGTTTACCT
TCACAGCTCCAAAAATGGGACATACAAATTTTGTCTCAGCAATGAATTTTCTACTTTCACACATAAACTGTATATTTTGATT
20 TTCAAGTTGGAGAGACCCACCTCTGTTTCTTAGTAGACCGAGTCAGTGCTCTTACCCAGATGGAATCTGCCTGTGTTCAATTC
ACGAAGCTCTGAAGTCTGTCATCGATTATCAGACTCATTTCCGTTTAAAGAGAAGCTCAAGGCCGAAGCCGAGCAGAGGATCTAA
ATACAAGAGTGGCCTATTGGTCAGTAGGAGAAGCCCTCATTCTTCTGGTGGTTAGCATAGGGCAGGTATTTCTTTTGAAGCT
TTTCTCAGATAAAAGAACACCACAACCTCGTGTGGATCATAACTACGTTTGTAGAATTGATGCACCATTGCCACTGTAATAT
TGCTGTCTCTAATTAATTTTAGGTACTGAAGAACTTAATATTGGCAACATTTTAAATCCTTACTCATACACTTGTGGGAGG
25 GATGTACAATGCATATTTCCAAACTGTGGAAGGACACCTTTTATTGTAAGGTGGAAGAACTTTGGAACATCATTTTGGG
CTATTATGTTAAATATTCAACCAATGATCTACTCTGTTGCGAGTTGTTATATCTACTCTTCGCACACTAACTTTGGTAT
TTTGATTCTTTTAAACATTTAAGACTACTTTTCTTATAGGTAGTTGATATTTAAAACTTTAGATTAAATGCTACATGTGT
TAGGGAGGAAGAAAAATAAATCTACCATAGGAATGTTAGTTAAATATTGATAATCAATTTTGAATAACAGTGGAGCTGAAAA
CAGTTTATGGAGGAAAAATAAATCTACCATAGGAATGTTAGTTAAATATTGATAATCAATTTTGAATAACAGTGGAGCTGAAAA
30 GCAATTTTATTCTTTTGAAGATTTCATTTAAATCTCTGTACTATGCATTTTGTGTTTTCAGTTTGTAAATGTTAATTTATAGT
TAAATAGGACAAAGCAGGCCCTATGACTATTGTTTCAAAAAATTAATATATAAATAAATTCAGAGGGGTTTTCATTTACTTTT
GAAAAGAAATTTCTTTAACAGGATTGCACCTATCAATTTCTTGTAGAATAAAAAAAATCTTTCAACAAA

291

MNGFTPDMSRGGDAAAAVAAVVAAAAAASAGNCTGAGTGAEVPGAGAVSAAGPPGAAGPGPQLCCLREDGERCCRAAGNAS
35 FSKRIQKSI SQKVKI ELDKSARHLYICDYHKNLIQSVNRNRKRKGSDDDGSDSPVQDIDTPEVDLYQLQVNTLRRYKRHFCLP
TRPGLNKAQLVEIVGCHFRSIPVNEKDTLYFIYSVKNDKNSDLKVDGSGVH

292

CCCCATGTGACAGTGACGGGGTCCCCGCTCCAGGAGACGCTCGAGTCTGCGTCCCGGCCCTCAGCACTGTCCACTGTTTCGGT
GCCAGCAGAGACCAGCAGGCCCGGGACAGTTGGTGTTTGGCCGTGCGCTGTCTAACTTGGTGTGCAGAGTGAATTCGGCTGC
40 CGGAGCGGAGAGAGCGGAGCGGCCAGGAGAGAGGGGATTTCTGTCAGCGCCGGCTCGGGAGCTCGGAGACATGAACGGCTTC
ACGCTTGACGAGATGAGCCGCGCGGGGATGCGGCCGCGCAGTGGCCGAGTGCTGCTGCGCGCGCGCGCGCGCTCGGCG
GGGAACGGGACCGCGCGCGGACCGGGCTGAGGTGCGCGCGCGGGGCGGTCTCAGCGGCTGGGCCCCGGGGCGCGCGG
CCGGGCCCCGGGCAACTGTGCTGCTGCGGGAGGATGCTGAGCGGTGCGGCGCGGCGGAGCAAGCCAGCTTCAGCAAGAGG
ATCCAGAAGAGCATCTCCAGAAGAAGGTGAAGATCGAGCTGGATAAGAGCGCAAGGCATCTTTACATATGTGATTATCATAAA
45 AACTTAATTCAGAGTGTTTCAAAACAGAAGAAAGAGAAAAGGGAGTGATGATGATGGAGGTGATTACCTGTTCAAGATATTGAT
ACCCAGAGGTTGATTTATACCAATTACAAGTAAATACACTTAGGAGATACAAAAGACACTTCAAGCTACCAACCAGACCAGGA
CTTAATAAAGCACAACCTGTTGAGATAGTTGGTTGCCACTTTAGGTCTATTCCAGTGAATGAAAAAGACACCTTAACATATTTT
ATCTACTCAGTGAAGAATGACAAGAACAATCAGATCTCAAGGTGATAGTGGTGTTCCTAGGAGACGTGGAATTGAGACTAA

TAACTTGGATGTTAACTACTGTTTACTGTTTTTTCACATGTAGAAATGTTCTTTGTGTATTTTTTCTACAGAGGATTTTCTCTGA
TTTTATTTTCTTTTGTCTGACTCTAATAATTAGTTGGAAATCATATAAAATGAGCTTTCCATAAATTAATCTATTTTAAATA
AAGGTTATTACTATTAAAAA

293

- 5 MPSAGTLPWVQGIICNANNPCFRYPPTGPEAPGVVGNFNKSIVARLFSDARRLLYSQKDTSMKDMRKVLRTLQQIKKSSSNLKL
QDFLVDNETFSGLYHNLSLPKSTVDKMLRADVILHKVFLQGYQLHLTSLCNGSKSEMIQLGQDEVELCGLPREKLAAAEV
LRNMDILKPILRTLNSTSPFPKELABATKTLHSLGLTAEQELFSMRWSMDMRQEVMTLNWNSSSSSTQIYQAVSRIVCGHP
EGGGLKIKSLNWDYEDNNYKALFGGNGTEDAETFYDNSTTPYCNLDLMKNLESSPLSRIWKALKPLLVGKILYTPDTPATRQVM
AEVNKTQELAVPHDLEGMWHEELSPKIWTFMENSQEMDLVRMLDSDRNDHFWEQQLDGLDWTQDQIVAFIAKHPEDVQSSNGS
10 VYTWREAFNETNQAIRTISRMECVNLNKLPIATEVWLINKSMELLDERKFWAGIVFTGITPGSIELPHHVYKIRMDIDNVE
RTNLIKDGWYDGPGRADPFEDMRYVWGGFAYLQDVVEQAIIRVLTGTEKKTGVYMQMPYPCYVDDIFLRVMSRSMPLFMTLAW
IYSVAVIIGKIVYEKEARLKETMRIMGLDNSILWFSWFISSLIPLLVSAAGLLVVLKGLNLLPYSDPSVVFVFLSVFAVVTILQ
CFLISTLFSRANLAAACGGIIFYTLVLPVLCVAVQDYVGTFLKIFASLLSPVAFGFGCEYFALFEEQIGVQWDLNLFESPVEE
DGFNLTSVSMMLFDTFLYGVMTWYIEAVFPQGYGIPRPWYFCTKSYWFGESDEKSHPGSNQKRISIEICMEEPTHLKLGVS
15 IQNLVKVYRDGMKVAVDGLALNFYEGQITSFLGHNGAGKTTTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQNLGVCPOHNV
FDMLTVEEHIWFIYARLKLSEKHVKAEMEQMALDVLPSKLSKTSQSLSGMQRKLSVALAFVGGSKVILDEPTAGVDPYSR
RGIWELLKYRQRTIILSTHMDADVLGDRIAIISHGKLCVSSFLKNQLGTGYLTLVKKDVSSLSRNSSTSVSYL
KKEDSVSQSSDAGLSDHESDTLTIDVSAISNLIRKHVSEARLVEDIGHELTYVLPYEAKEGAFVELFHEIDRLSDLGIS
YGISETTLEEIFLKVAAESGVDAETSDGTLPARNRRAFGDKQSCLRPFTEDDAADPNDSIDIPESRFTDLGMDGKGSYQVK
20 GWKLTQQQFVALWKRLLIARRSRKGFQAQIVLPAVFCIALVPSLIVPPFGKYPSLELQPMWYNEQYTFVSNDAPEDTGTLEL
LNLTKDPCFGTRCMEGNPIPDTPCQAGEEWTAPVPQTIMDLFQNGNWTMQNPSPACQCSSDKIKMLPVCPPGAGGLPPPQ
RKQNTADILQDLTGRNISDYLKTYVQIIAKSLKNKIWNNEFRYGGFSLGVSNTQALPPSQEVNDATQMKHLKLAKDSSADR
FLNSLGRFMTGLDTRNNVWVFNKNGWHAISSFLNINNAILRANLQKGENPSHYGITAFNHPNLTKQLSEVAPMTTSVDVL
VSICVIFAMSFVPASFVFLIQUERVS KAKHLQPISGVKPIYWLNSFVWDMCNVVPATLVIIIFICFQKSYVSSSTNLPVLAL
25 LLLLYGWSITPLMPASFVFKIPSTAYVLTSVNLFIGINGSVATFVLELFTDNKLNINDILKSVFLIFPHFCLGRGLIDMV
NQAMADALERFGENRFVSPSLWDLVGRNLFAMAVEGVVFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQRIIDGGGQ
NDILEIKELTKIYRRKRKPAVDRIKCVGIPPEGCFGLLVNGAGKSSFTKMLTGDTTIVTRGDAFLNRNSILSNIHEVHQNMGYC
QFDAITELLTGREHVEFFALLRGVPEKEVGKVGWEAIRKGLVYKEGYAGNYSNGNKRKLSTAMALIGPPVVFVFLDEPTTGMD
PKARRFLWNCALSVVKEGRSVVLTSHSMEECEALCTMAIMVNGRFRCLGSVQHLKNRFGDGYTIVVRIAGSNPDLKPVQDFFG
30 LAFPGSVPEKHRNMLQYQLPSSLSSLARIFSISSQSKRLHIEDYSVSQTTLDQVFNFAKDQSDDDLKDLSLHKNQTVVDV
AVLTSFLQDEKVKESYV

294

- CAACATGTCTAGCTGTTACTGGAAGTGGCCTGGCCTCTATTTATCTTCTCTGATCTCTGTTTCGGCTGAGCTACCCACCC
TATGAACAACATGAATGCCATTTTCCAAATAAGCCATGCCCTCTGCAGGAACACTTCTTGGGTTTCAGGGGATTATCTGTAAAT
35 GCCAACAACCCCTGTTTCCGTTACCCGACTCCTGGGAGGCTCCCGAGTGTGTGGAACTTTAACAAATCCATTGTGGCTCGC
CTGTTCTCAGATGCTCGGAGGCTTCTTTTATACAGCCAGAAAGACACCAGCATGAAGGACATGCGCAAAGTTCTGAGAACATTA
CAGCAGATCAAGAAATCCAGCTCAAACCTGAAGCTTCAAGATTTCTTGGTGACAATGAAACCTTCTCTGGGTTCTGTATCAC
AACCTCTCTCTCCCAAAGTCTACTGTGGACAAGATGCTGAGGGCTGATGTCATTCTCCACAAGGTATTTTGAAGGCTACAG
TTACATTTGACAAGTCTGTGCAATGGATCAAATCAGAAGAGATGATTCACCTTGGTGACCAAGAAGTTCTGAGCTTTGTGGC
40 CTACCAAGGAGAACTGGCTGCAGCAGAGCGAGTACTTCGTTCCAACATGGACATCCTGAAGCAATCCTGAGAACATAAAC
TCTACATCTCCCTTCCGAGCAAGGAGCTGGCCGAAGCCACAAAACATTCGTGCATAGTCTTGGGACTCTGGCCAGGAGCTG
TTCAGCATGAGAAGCTGGAGTGACATGCGACAGGAGGTGATGTTCTTQACCAATGTGAACAGCTCCAGCTCCTCCACCAATC
TACCAGGCTGTGTCTCGTATTGTCTCGCGGCATCCCGAGGAGGGGGCTGAAGATCAAGTCTCTCACTGGTATGAGGACAAC
AACTACAAAGCCCTCTTTGGAGGCAATGGCACTGAGGAAGATGCTGAAACCTTCTATGACAACCTCTACAACTCTTACTGCAAT
45 GATTTGATGAAGAATTTGGAGTCTAGTCTCTTTTCCCGATTATCTGGAAGCTCTGAAGCCGCTGCTCGTTGGGAAGATCCTG
TATACACCTGACACTCCAGCCACAAGGCAGGTCTAGGCTGAGGTGAACAAGACCTTCCAGGAAGTGGCTGTGTTCCATGATCTG
GAAGGCATGTGGGAGGAACCTCAGCCCAAGATCTGGACCTTCATGGAGAACAGCAAGAAATGGACCTTGTCCGGATGCTGTTG
GACAGCAGGACAAATGACCACTTTTGGGAACAGCAGTTGGATGGCTTAGATTGGACAGCCCAAGACATCGTGGCGTTTTTGGCC
AAGACCCAGAGGATGTCCAGTCCAGTAATGGTTCTGTGTACCTGGAGAGAAGCTTTCAACGAGACTAACAGGCAATCCGG
50 ACCATATCTCGCTTCATGGAGTGTGTCAACCTGAACAAGCTAGAACCCATAGCAACAGAAGTCTGGCTCATCAACAAGTCCATG

GAGCTGCTGGATGAGAGGAAGTTCTGGGCTGGTATTGTGTTTACTGGAATTACTCCAGGCAGCATTGAGCTGCCCCATCATGTC
AAGTACAAGATCCGAATGGACATGACAAATGTGGAGAGGACAAATAAAATCAAGGATGGGTACTGGGACCTTGGTCTCGAGCT
GACCCCTTTGAGGACATGCGGTACGTCTGGGGGGCTTCGCCTACTTGCAGGATGTGGTGGAGCAGGCAATCATCAGGGTGCTG
ACGGGCACCGAGAAGAAAAGTGGTGTCTATATGCAACAGATGCCCTATCCCTGTACGTTGATGACATCTTCTGCGGGTGATG
5 AGCCGGTCAATGCCCTCTTCATGACGCTGGCCTGGATTACTCAGTGGCTGTGATCATCAAGGCGATCGTGTATGAGAAGGAG
GCACGGCTGAAAGAGACCATGCGGATCATGGGCTGGACAACAGCATCCTCTGGTTTAGCTGGTTTCAATAGTAGCCTCATTCCT
CTTCTGTGAGCGCTGGCCTGCTAGTGGTCATCCTGAAGTTAGGAAACCTGCTGCCCTACAGTGATCCCAGCGTGGTGTGTC
TTCTGTCCGTGTTTGTGTGGTGACAATCCTGCAGTGCTTCTGTATTAGCACACTTCTTCCAGAGCCAACCTGGCAGCAGCC
TGTTGGGGCATCATCTACTTCACGCTGTACCTGCCCTACGTCTGTGTGTTGGCATGGCAGGACTACGTGGGCTTCACACTCAAG
10 ATCTTCGCTAGCCTGCTCTCTCTGTGGCTTTTGGGTTTGGCTGTGAGTACTTTGCCCTTTTGGAGGAGCAGGGCATTGGAGTG
CAGTGGGACAACCTGTTTGGAGTCTCTGTGGAGGAAGATGGCTTCAATCTCACCCTTCGGTCTCCATGATGCTGTTTGACACC
TTCTCTATGGGGTGATGACCTGGTACATTGAGGCTGTCTTCCAGGCCAGTACGGAATTTCCAGGCCCTGGTATTTTCTCTG
ACCAAGTCTACTGGTTTGGCGAGGAAAGTGATGAGAAGAGCCACCTGGTTCCAACAGAGAAGAAATCAGAAATCTGCATG
GAGGAGGAACCCACCCACTTGAAGCTGGGCGTGTCCATTGAGAACCTGGTAAAAGTCTACCGAGATGGGATGAAGGTGGCTGTC
15 GATGGCCTGGCACTGAATTTTATGAGGGCCAGATCACCTCCTTCTGGGCCACAATGGAGCGGGGAAGACGACCACCATGTCA
ATCTGACCGGGTGTTCCTCCCGACCTCGGGCACCGCTACATCTGGGAAAAGACATTTCGCTCTGAGATGAGCACCATCCGG
CAGAACCTGGGGTCTGTCCCGACATAACGTGCTGTTTGCATGCTGACTGTGCAAGAACACATCTGGTTCTATGCCCGCTTG
AAAGGGCTCTTGAGAAGCAGTGAAGGCGGAGATGGAGCAGATGGCCCTGGATGTTGGTTTGGCATCAAGCAAGCTGAAAAGC
AAAACAAGCCAGCTGTGAGGTGAATGCAGAGAAAGCTATCTGTGGCTTGGCCCTTGTGCGGGGATCTAAGGTTGTCTATCTG
20 GATGAACCCACAGCTGGTGTGGACCTTACTCCCGCAGGGGAATATGGAGCTGCTGCTGAAATACCGACAAGGCCGACCAT
ATTCTCTTACACACCACATGGATGAAGCGGACGCTCTGGGGACAGGATTGCCATCATCTCCCATGGGAAGCTGTGCTGTGTG
GGCTCTCTCTGTTCTGTAAGAACCAGCTGGGAACAGGCTACTACCTGACCTTGGTCAAGAAAGATGTGGAATCTCCCTCAGT
TCCTGCAGAAACAGTAGTAGCACTGTGTATACCTGAAAAGGAGGACAGTGTCTCTCAGAGCAGTCTCTGATGCTGGCCTGGG
AGCGACCATGAGAGTGACACGCTGACCATCGATGCTCTGCTATCTCCAACCTCATCAGGAAGCATGTGTCTGAAGCCCGCTG
25 GTGGAAGACATAGGGCATGAGCTGACCTATGTGCTGCCATATGAAGCTGCTAAGGAGGGAGCCTTTGTGGAACCTTTTCATGAG
ATTGATGACCGCTCTCAGACCTGGGCATTTCTAGTTATGGCATCTCAGAGACGACCTTGAAGAAATATTTCTCAAGGTGGCC
GAAGAGAGTGGGGTGGATGCTGAGACCTCAGATGGTACCTTGGCAGCAAGACGAAACAGGCGGGCTTCGGGGACAAGCAGAGC
TGCTCTCGCCGTTCACTGAAGATGATGCTGCTGATCCAATGATTTCTGACATAGACCCAGAATCCAGAGAGACAGACTTGCTC
AGTGGGATGGATGGCAAGGGTCTTACCAGGTGAAAGGCTGGAACCTTACACAGCAACAGTTTGTGGCCCTTTTGTGGAAGAGA
30 CTGCTAAATGCCCAGACGGAGTCCGAAAGGATTTTGTGCTCAGATTGTCTTGGCAGCTGTGTTTGTCTGATTGCCCTTGTGTTT
AGCCTGATCGTGCCACCCCTTGGCAAGTACCCAGCCTGGAACCTTACGCCCTGGATGTACAACGAAACAGTACACATTTGTGAGC
AATGATGCTCTGAGGACACGGGAACCTTGAACCTTAAACGCCCTCACCAAGACCCCTGGCTTCGGGACCCGCTGTATGGAA
GGAAACCCAAATCCAGACACGCCCTGCCAGGAGGGGAGGAAGAGTGGACCACTGCCCCAGTTCCCCAGACCATCATGGACCTC
TTCCAGAATGGGAAGTGGACAATGCAGAACCCTTACCTGCATGCCAGTGTAGCAGCGACAAAATCAAGAAGATGCTGCCTGTG
35 TGTCCCCAGGGGAGGGGGGCTGCCTCTCCACAAAGAAAACAAACACTGCAGATATCCTTCAGGACCTGACAGGAAGAAAC
ATTTTCGATTATCTGGTGAAGACGTATGTGCAGATCATAGCCAAAAGCTTAAAGAACAAGATCTGGGTGAATGAGTTAGGTAT
GGCGGCTTTTCCCTGGGTGTCAGTAATACTCAAGCACTTCTCCAGTCAAGAAGTTAATGATGCCACCAACAAATGAAGAAA
CACCTAAAGCTGGCCAAGGACAGTTCTGCAGATCGATTTCTCAACAGCTTGGGAAGATTTATGACAGGACTGGACACCAGAAAT
AATGTCAAGGTGTGTTTCAATAACAAGGGCTGGCATGCAATCAGCTCTTCTCTGAATGTATCAACAATGCCATTTCTCCGGGCC
40 AACCTGCAAAAGGGAGAGAACCTTAGCCATTATGGAATTACTGCTTTCAATCATCCCTGAATCTCACCAAGCAGCAGCTCTCA
GAGGTGGCTCCGATGACCACATCAGTGGATGTCTTGTGTCATCTGTGTCATCTTTGCAATGTCTTCTGCCAGCCAGCTTT
GTCTGATTTCTGATCCAGGAGCGGGTCAGCAAAGCAAAACACCTGCAGTTTATCATCTTCTATCTGCTTCCAGCAGAAGTCC
TCTAATTTTGTCTGGGATATGTGCAATTACGTTGTCCCTGCCACACTGGTCAATTATCATCTTCTATCTGCTTCCAGCAGAAGTCC
TATGTGCTCTCCACCAATCTGCCTGTGCTAGCCCTTCTACTTTTGTCTGTATGGGTGGTCAATCACACCTCTCATGTACCCAGCC
45 TCCTTTGTGTTTCAAGATCCCCAGCACAGCCTATGTGGTGCTCACCAGCGTGAACCTCTTCAATTGGCATTAAATGGCAGCGTGGCC
ACCCTTGTGTGGAGCTGTTTCAACGACAATAAGCTGAATAATATCAATGATATCCTGAAGTCCGTGTTCTTGATCTTCCACAT
TTTGTCCCTGGGACGAGGGCTCATCGACATGGTGAAAAACAGGCAATGGCTGATGCCCTGGAAAGGTTTGGGGAGAATCGCTTT
GTGTACCAATTATCTTGGGACTTGGTGGGACGAAACCTCTTCGCCATGGCGTGGAGGGGTGGTGTCTTCTCTCATTTCTGTT
CTGATCCAGTACAGATTCTTTCATCAGGCCAGACCTGTAATGCAAGCTATCTCTCTGAATGATGAAGATGAAGATGTGAGG
50 CGGGAAAGACAGAGAATTTGATGGTGGAGGCCAGAATGACATCTTAGAAATCAAGGAGTTGACGAAGATATATAGAAGGAAG
CGGAAGCCTGCTGTGACAGGATTTGCGTGGGCAATTCCTCTGGTGAGTGTCTTGGGCTCTGGGAGTTAATGGGGCTGGAAAA
TCATCAACTTTCAAGATGTTAACAGGAGATACCCTGTTTACCAGAGGAGATGCTTTCTTAAACAGAAATAGTATCTTATCAAAAC

ATCCATGAAGTACATCAGAACATGGGCTACTGCCCTCAGTTTGTATGCCATCACAGAGCTGTTGACTGGGAGAGAACACCTGGAG
TTCTTTGCCCTTTTGGAGAGGAGTCCCAGAGAAAGAGTTGGCAAGGTTGGTGAGTGGGCGATTCCGAAACTGGGCCCTCGTGAAG
TATGGAGAAAAATATGCTGGTAACATATAGTGGAGGCAACAAACGCAAGCTCTCTACAGCCATGGCTTTGATCGGCGGGCCTCCT
GTGGTGTCTTCTGGATGAACCCACCACAGGCATGGATCCCAAAGCCCGCGGTTCTTGTGGAATTGTGCCCTAAGTGTGTCAAG
5 GAGGGGAGATCAGTAGTCTTACATCTCATAGTATGGAAGAATGTGAAGCTCTTTCAGTACTAGGATGGCAATCATGGTCAATGGA
AGGTTTCAGGTGCCCTTGGCAGTGTCCAGCATCTAAAAAATAGGTTTGGAGATGGTTATACAATAGTTGTACGAATAGCAGGGTCC
AACCCGGACCTGAAGCCTGTCCAGGATTTCTTTGGACTTGCATTTCCCTGGAAGTGTTCAAAAGAGAAACACCGGAACATGCTA
CAATACCAGCTTCCATCTTCATTTATCTTCTCTGCGCCAGGATATTCAGCATCCTCTCCCAGAGCAAAAAGCGACTCCACATAGAA
GACTACTCTGTTTCTCAGACAACACTTGACCAAGTATTTGTGAACCTTGCCAAGGACCAAAAGTGATGATGACCATTAAAAAGAC
10 CTCTCATTACACAAAAACCAGACAGTAGTGGACGTTGCAGTTCTCAGATCTTTTCTACAGGATGAGAAAGTGAAGAAAGCTAT
GTATGAAGAATCCTGTTTATACAGGGGTGGCTGAAAGTAAAGAGGNACTAGACTTTCCTTTCAGCCATGTGAAGTGTGTGGAGA
AAAGAGCCAGAAGTTGATGTGGGAAGAAGTAACTGGATACTGTACTGATACTATTCAATGCAATGCAATTCATGCAATGAA
ACAAAATTCATTACAGGGGCAGTGCCTTTGTAGCCTATGTCTTGTATGGCTCTCAAGTGAAAGACTTGAATTTAGTTTTTTTAC
CTATACCTATGTGAAACTCTATTATGGAACCCAAATGGACATATGGGTTTGAATCAGCTTTTCTTTTCTTTTGTCTCTGTGT
15 ATTCTCATTGGGGTTGCAACAATAATTCATCAAGTAATCATGGCCAGCGATTATTGATCAAAATCAAAGGTAATGCACATCCT
CATTCATAAGCCATGCCATGCCAGGAGACTGGTTTCCCGGTGACACATCCATTGCTGGCAATGAGTGTGCCAGAGTTATTAG
TGCCAAGTTTTTTCAGAAAGTTTGAAGCACCATGGTGTGTCATGCTCAGCTTTTGTGAAAGCTGCTCTGCTCAGAGTCTATCAACA
TTGAATATCAGTTGACAGAATGGTGCCATGCGTGCTTAACATCCTGCTTTGATTCCCTCTGATAAGCTGTCTGCTGGCAGTAA
CATGCAACAAAAATGTGGGTGTCTTAGGCACGGGAAACTTGGTTCCATTGTATTATTTGCTTATGCTTCGAGCCATGGGTCTA
20 CAGGGTCATCCTTATGAGACTCTTAAATATACTTAGATCCTGGTAAAGAGGCAAGAATCAACAGCCAAACTGCTGGGGCTGCAA
GCTGCTGAAGCCAGGGCATGGGATTAAGAGATTGTGCGTTCAACCTAGGGAAGCCTGTGCCCATTTGTCTGACTGTCTGCT
AACATGGTACACTGCATCTCAAGATGTTTATCTGACACAAGTGATTATTTCTGGCTTTTGAATTAATCTAGAAAATGAAAG
ATGGAGTTGTATTTTGACAAAAATGTTTGTACTTTTAAATGTTATTTGGAATTTAAGTTCTATCAGTGACTTCTGAATCCTTA
GAATGGCCTCTTTGTAGAACCCGTGTGGTATAGAGGAGTATGGCCACTGCCCCACTATTTTATTTTCTTATGTAAGTTTGCATA
25 TCAGTCATGACTAGTGCCTAGAAAGCAATGTGATGGTCAGGATCTCATGACATTATATTTGAGTTTCTTTTCAGATCATTTAGGA
TACTCTTAATCTCACTTCATCAATCAAAATATTTTGTAGTGTATGCTGTAGCTGAAAGAGTATGTACGTACGTATAAGACTAGA
GAGATATTAAAGTCTCAGTACACTTCTGTGCCATGTTATTCAGCTCAGCTGGTTTACAAATATAGGTTGTCTTGTGGTTGTAGGA
GCCCCACTGTAAACAATACTGGGCAGCCTTTTCTTTTCTTTTAAATGCAACAATGCAAAAGCCAAGAAAGTATAAGGGTCACAAG
TCTAAACAATGAATTCTTCAACAGGGAAAAACAGCTAGCTTGAAAACTTGTGAAAAACACAACCTTGTGTTTATGGCATTTAGTA
30 CCTTCAAAATAATTGGCTTTGCAGATATTGGATACCCCAATAATCTGACAGTCTCAAAATTTTTCATCTCTTCAATCACTAGTCA
AGAAAAATATAAAAAACAACAATACTTCCATATGGAGCATTTTTCAGAGTTTCTAACCAGTCTTATTTTCTAGTCAGTAAA
CATTTGTAAAAATACTGTTTCACTAATACTTACTGTTAACTGTCTTGAGAGAAAAAGAAAAATATGAGAGAACTATTGTTTGGG
AAGTTCAAGTGATCTTTCAATATCATTACTAACTTCTTCCACTTTTTCAAAATTTGAATATTAACGCTAAAGGTGAAGACTT
CAGATTTCAAATTAATCTTCTATATTTTAAATTTACAGAATATATATAACCCACTGCTGAAAAAGAAAAAATGATTGTT
35 TTAGAAGTTAAAGTCAATATGATTTTAAATATAAGTAATGAAGGCATATTTCCAATAACTAGTGATATGCGATCGTTGCATTT
TACAGTATCTTCAAAAATACAGAATTTATAGAATAATTTCTCCTCATTTAATATTTTCAAAAATCAAAGTTATGTTTCTCAT
TTTACTAAAAATCGTATTTCTAATTTTCATTATAGTAAATCTATGAGCAACTCCTTACTTCGGTTCTCTGATTTCAAGGCCATA
TTTTAAAAAATCAAAAGGCACTGTGAACTATTTTGAAGAAAAACACAACATTTTAAATACAGATTGAAGGACCTCTTCTGAAGCT
AGAAACAATCTATAGTTATACATCTTCATTAATACTGTGTTACCTTTTAAATAGTAATTTTTCATTTTCTCTGTGTAACCT
40 AATTGTGGTAGAAATTTTACCAACTCTATACTCAATCAAGCAAAATTTCTGTATATTCCTGTGGAAATGTACCTATGTGAGTT
TCAGAAATTTCAAAATACGTGTTCAAAAATTTCTGCTTTTGCATCTTTGGGACACCTCAGAAAACTTATTAACAACCTGTAAT
ATGAGAAATACAGAAGAAAAATAAGCCCTCTATACATAAATGCCAGCAAAATTCATTGTTAAAAACAACCAACCTCACA
CTACTGTATTTTCATTATCTGTACTGAAAGCAAATGCTTTGTGACTATTAAATGTTGCACATCATTTCATTTCAAAAAA
AAAAA
45 295
PPGPERSRLGLVSLHQRSCPKICIAVFTRVSEPRIQFPASRILPSSNTSKDFDPVSGQSNYGSQSGQTLNRPVVASNPVTPS
LHSGPAPRMPPLPASQNPATTPMPSSFLPEANLPPPLNWQYNYNSTASQTNHCPRASSQPTVSGNTSLTNNHQYVSSGYPSLQN
SFIKSGPSVPLVNPPLPTTFQPGAPHGPPAGGPPVRLTPLTSSYRDVPQPLFNSAVNQEGITSNTNNGSMVHVSSYDEIE
GGGLLATPQLTNKNPKMSRSVGYSPSLPPGYQNTTPPGATGVPSSLNYPSPGPQFTQTPLGANHLTSSMGLSLQPEGLRVV
50 NLLQERNMLPSTPLKPPVNLHEDIQKLNCNPELFRCTLSIPQTOALLNKAKLPLGLLLHFPKDLVQLPVVTSSTIVRCRSCR
TYINPFPVSFLDQRRWKCNCYRVNDVPEEFLYNPLTRVYGEFHRREPVQNTIEFMAPSEYMLRPPQPPVYLFVFDVSHNAVET

- GYLNSVCQSLLDNLDPGNTRTKIGFITFDSTIHFYGLQESLSQPQMLIVSDIEDVFI MPENLLVNLNESKELVQDLLKTL P
QMFTKTLETQSGALGPALQAAFKLMSPTGGRMSVFQTLPLTGVGALKPREPNHRSSAKDIHMTPTDFYKKLALDCSGQQVAV
DLFLLSGQYSDLASLGCISRY SAGSVYYPYHHQHNPVQVQKLQKELQRYLTRKIGFEAVMRIRCTKGLSIHTFHCFNFFVRST
DLLSLPNVNPADAGYAVQMSVEESLTDQLVSFQ SALLYTSSKGERRIRVHTLCLPVVSTLNDVFLGADVQAISGLLANMAVDRS
5 MTASLSDARDALVNAVIDSL SAYRSSVLSNQPGMLVPFSLRLFPLFVLALLKQKSFQGTGNARLDERIFAMCQVKNQPLVYLM
LTHPSLYRVDNLSDGALNISDRITIPPILQLSVEKLSRDGAFLMDAGSVLMLWVGKNCTQNFLSQVLGVQNYASI PQPMTD
LPELDTPE SARIIAFISWLRQRPFFPILYVIRDESPMKANFLQNMIEDRTESALSYYEFLHLHIQQQVNK . .
- 296
- 10 CCTCCAGGCCCAGAACGGAGCCGCTTGGCCTCGGGGTCTCCCTACACCAACGGTCTGTCCAAAATGCATTGCTGTCTTCACA
AGAGTCAGTGAGCCAGGATACAATTTCAGCTTCCAGGATCTACCCCTCATCCAATACCAGCAAAGACTTTGATCCAGTCTCT
GGCCAGTCTAACTATGGTGGTCTCAGGGATCTGGGCAGACTCTTAATAGACCACCTGTGGCCTCTAATCCAGTGACACCTTCG
CTTCATAGTGGTCTGTCTCCCGAATGCCATTACCTGCTTCTCAGAACCCAGCTACTACACCAATGCCTTCTAGTAGCTTTCTT
CCTGAAGCCAACTGCCACCACCTTTGAATTGGCAATATACTATCCATCCACAGCCTCACAACAAACCATTGTCTCTGTGCA
TCATCCCAACCACTGTATCTGGAATAACAAGTTTAACCACAAATCATCAATATGTTTCTTCTGGATATCTTTCATTCAAAT
15 AGCTTCATAAAGTCAGGTCTTCTGTACCTCCCTTAGTGAATCCACCTCTGCCTACAACCTTTCAACCAGGAGCTCTCATGGG
CCCCCTCCAGCTGGAGGCCACCCCGAGTGAGGGCCCTCAGCCCCCTGACATCATATAGAGATGTACCCAGCCCTTATTT
AATTCAGCTGTCAACCAAGAAGGTATTACATCAAATACCAATAACGGATCTATGGTGGTCCACAGTAGTTACGACGAGATTGAA
GGAGGTGGCTTATTGGCAACACCACAGCTTACTAACAAGAAATCCCAAAATGAGCCGAAGTGTGGATATTATATCCCTCCTTA
CCACCTGGTTATCAGAACACACACCACCTGGTGCAACTGGAGTACCACCTCTTCTTGAATTACCCAAGTGGGCCACAAGCC
20 TTTACTCAGACTCCCTTAGGTGCTAATCATTAAACCACAAGCATGAGTGGATTAAGTCTACAACCAGAGGGTCTAAGAGTTGTC
AATCTTCTTCAAGAAAGAAACATGCTTCCGTCAACACCTTTGAAGCCTCCAGTTCCAAATTTGCATGAAGACATCCAGAACTC
AACTGTAACCCAGAGTTATTTTCGATGCACGCTGACTAGCATTCCTCAGACGAGGCCCTTATTGAATAAAGCCAACTTCTTTG
GGGCTGCTGCTTCATCTTCAAAGACTTAGTGCAATTGCTGTGGTTACCTCCAGTACAATTTGTGAGATGCCGTTTCATGCAGG
ACGTACATCAATCTTTCGTGAGCTTTCTTGATCAAAGGAGATGGAAGTGAACCTTATGTTATCGAGTCAATGATGTTCTTGAA
25 GAATTCCTTGTAACAACCTTTGACCAGAGTTTATGGAGAACCTCAGAGAAGCCAGAAAGTTCAAAATGCTACTATTGAGTTTATG
GCTCCTTCAGAAATACATGTTACGACCACCTCAGCCTCAGTGTATCTCTTGTATTGATGTGCTCACAATGCAGTCGAAACT
GGATACTTGAATTCAGTTTGCCAGAGTTTGTAGACAATCTGGATTGCTTCTGCGCAACACTAGAACAAAAATTGGCTTCATA
ACATTTGACAGTACAATCCATTTCTACGGTCTTCAGGAAAGTCTCTCTCAACCTCAGATGCTAATAGTTTCAGATATTGAAGAT
GTTTTTATACCTATGCCAGAGAACTTATTAGTAAACTTAAATGAAAGTAAAGAGCTCGTGCAAGATTTACTGAAAACCTTGCCA
30 CAAATGTTTACCAAGACTCTGGAGACCCAGAGTGCCTTGGGCTCTGCACTGCAGGCTGCCTTTAAGCTGATGTCTCCAACCTGGT
GGTCGAATGTCTGTCTTCAAACACAACCTCCCAACTCTTGGAGTGGGAGCCCTGAAACCACGAGAGGAACCAACCACAGGTCA
TCTGCTAAGGATATACACATGACACCATCCACTGACTTCTATAAGAAATTAGCCTTGGACTGTTCTGGTCAGCAAGTTGCTGTT
GACTTATCTCTCTCAGTGGACAGTATTCTGATTTGGCTTCTCTGGGTTGTATTTCTCGGTATTGACAGGAGTGTCTATTAC
TATCCCTCTTACCATCATCAGCACAACCCAGTCCAAGTACAGAAATTACAGAGGAACCTACAGAGATACCTTACTCGGAAGATT
35 GGCTTTGAGGCAGTCATGAGGATTCGGTGCAACAAAGTCTTTCATTCATACTTTCCATGGAACCTCTTTGTTAGGTCAACC
GACTTACTGTCTTTGCTTAACGTCAACCCAGACGCTGGGTATGCAATGATGTCAGTGGAGAGAGTCTTACTGACACTCAG
TTGGTTTCTTTTCAGTCAGCACTCTTGATATACATCCAGCAAAGCGCAAGAAAGAAATTCGTGTTTCATACTTTGTGTTTGGCAGTA
GTTTCGACTCTGAATGATGTCTTCTTGGAGCTGATGTTCAAGCAATTTAGGGTTATTGGCCAATATGGCTGTTGACAGATCT
ATGACTGCCAGTCTGAGTGACGCTCGGGATGCTCTAGTGAATGCAGTCATTGACTCCCTTTTCAGCTTACCGTTCTTCAGTCTTA
40 AGTAACCAGCAGCCTGGAATCATGGTTCCTTTTCTTTGGGCTTTTCCCACTTTTGTGTTGGCTCTCCTTAAACAGAAATCA
TTCCAGACTGGGACAAATGCACGCTTAGATGAACGCAATTTTGTCTATGTGTCAAGTGAACCAACAGCCCTTGGTTTACCTTATG
CTCACAACTCATCCAGTTTGTATAGAGTTGACAATCTCTCAGATGAGGGAGCACTCAACATCAGTGATAGAACCATACCTCAG
CCCCCATCTTTCAGCTTTTCAGTGGAGAAGCTGAGCAGAGATGGAGCTTCTCTCATGGATGCAGGCTCTGTACTGATGCTTTGG
GTTGGAAAAAATTGTACACAGAAATTTCTCAGCCAAGTTCTAGGAGTTCAAAACTATGCATCAATTCACAGCCTATGACAGAC
45 CTTCCAGAACTTGATACACCAGAATCTGCCAGAATAATAGCTTTTCATCTCTTGGCTTAGAGAGCAGAGACCAATTTTCCCAATA
CTTTATGTAATAAGGGATGAGAGTCCAATGAAAGCAAACTTCTTCAAAACATGATAGAAGACAGAACAGAACTGCAATTATCA
TATTATGAATCTCTGTGCATATACAGCAACAAGTGAATAAATGAATGAATGAAGAAATTTGACTTATTTTTAAGGAATGTCAC
AATAGTGAGAAATACCTGGAATGTGTAATACCTTCTTTTTCTATTATGTTTGTGGACTAATGTGATGATTGAGATGTTCTCAC
TGTGATTTCAACAACCTATAAGCAAATAAAGACCACAGCAGAGAATCAAACATGCAACTCTGAAATACTGTATTTTCAAATC
50 AGAATATAGCTACGTATGATTGGATACCTTTTCTTGGCAATTATGTTGAGTTGTTATGGATTAAAAAAGAATATTGCAGAG
GCAAAGTACATTTGTAAAAATAAGATTCTGTGTTCTACATGTATTTCTCATTTTAAATTTTTCGAATCTTTGGCTGCTA

CATTTAAAACCTCACAAACCTAAGTGTTCAGGGAAGTTACAAATTGATTGGTAGTGATGTTTTTAAAAATAAAAACAATGGAAA
GTAAATATAATGTAGGAAAACCTAGAAATTCATTCCCCACACGTGCTTTTTTTTTTCTTTGTTAAGGGAAGGATCATGTTGACT
AAAATAAACTAAATTTCTAGTATAGCTTGAGAAAAATATGAAGAAACACATTCAGCTTTAAAAATCTGTCTAGTATTTCTTTATGCT
TGAAGAACAAAGTCACCTTTGATTGAAAGTGAGAACTATCATTAGGTGGTTTTCTGATTTCTTGATGAAGAGTTGGACATACTGT
5 CTTAATCTATAGTGAAGAATTTGAGCTGTCTTCATAAACACTGGGACTAGCAATGATAATAGGGAGATAAGAACTTTAATT
ATCTTGATCCTTTAAGTGGATTTTATTTGGTGCATTTCTGCTCTGGGTATATAATAAAGTGGGGTTTTTGGTGAAATGAGTG
AAGAAATGAAAGGTTTCTAAAGTGTCTATCCAAATACATCAGTAACATTTTCTAAGGAGTTTAATTGTTAAATTTGGAAGTCATT
CATAAGAAATATTTATGCTTGAATATGAAATCTATGAAAGCATAAATGCTGCTGTTGATTGGTGGATATTAAGATTATACA
CATCCAACATATTTAAAGTTATGAAAGAACTTGACTTCTGAAAAATCCTTAAGAGACTGCTTTCTTGATTTCAGCTAGAGAAATAT
10 TATAGTCAAAACTATTGAGTGAATTTTGTTTTACAAATAGGTAATTTATACATTTGTATATTTAAAGTGTCTGTGACATAGTATCT
TTAAGAGTTTGGCTCAGTTTTCACAGATTCATTTTGTCTTAAGAATTTCTTAAATATGTTTCATGTATAATACTTGATCAAAATA
TTTTTGGGTTTTTGTGTTTTTAAATGGGTTAGAAAATGTTTACAATCTTGGTCTTATATGATCACCATGGAATAGTAACCT
CCAGGTTTATATCAATATGAGCTGACTTTAACTGAGTTGTTGGGATAGGGAAGAGCAGTCCCTCTACAGTATACAACACTACTG
CTTGCCAGCTGGATCAAAATAATCATGTTTATGAAAAATATCTCCCTTAAGCAGTGTTAAGGTTGGTTTGCAGTGTGTAAGTGG
15 CACATTGAACTGGAGTTTTCTTGAAAGCTGCTTCATCTATTAAGAAGCAATTTTCAAATTTGTAGCGAATTTATATATCCCTC
TTTTAAAGAAACAGTCGTTATATGCTGATGTTTCTTAAATAACTAAAAATGTTCTCTTAAATGTGATTTTAAATGGAGTTATTT
GTAGGTCTTTCTTAGTAGTAAAGAATCTTCTAGAGGGAACATTTGTGCTTTTAGGGATAATCTTCTTGTGCTCTACATACAT
CCCTAAGTGGGTATGACTCTGTTATTACCACATGCTTTTTTAGTATATTTTCAAAATTTACTTTTAAATATTATTTAGATAC
GGTGAACATGTGCAATTCAGAATAATTTTATAACAGGTCATGAAAAACATAACTTTAGTTAGGATTCACAATATTGTTCTCC
20 ACATAATGAGAGATGAATGAGCCTTTGGAGATACGATATAAGGCAATTTATTTTGAATGTTGAATGTGTTTTTGTGTTGA
TTCTTTTTTTTTCCCCCAATAGGGCACTACCTGCCATATCATCTGTATTACTTTTTGATGTAAAGCGACTAATATTTACACTA
TGCCATATTTTTTTAATTATAGTTGTAAATTTATGAAAGATCCTTGAATTTTCTACAGATCTACAACACTAATGTAAACAGACA
AGGGCAATCTTGGTATTTAAATCTGAGCATGGCAGTTCTACCATAAAAAGTACTCTATTTTCTAATTTCTAGGATTTTTTAAAA
TAACATTTCTGTAAGTCTGACATACTAATAGTCACCAAGCAGTACCATTATTTTAGTTTGCATATATTTTCACTGTTTTTAA
25 TTTAATGTATGAGTCTAATAGACTGTTTTGCAATAATTAGAATAAAGATTTATTTCTTCTAATCAAGATGCATAACAGCTAT
TATCTAGGGGACCCTAAATGTGATTTCAAAATTTTGTAACTATTACAAATGTAATCCTTATATAGAAATTTTAAATTTGTGA
AAGTAGTGATAATATTGTAATATTAAATTTCTGTTCTTAAATTTCAAAATATGTATTGATCTTCAATGTGCTGTGTTAAATCTTG
CTTCTCTGAAAAGTTGGAGACAAGATTTGTCTTCTTTTTACAGTTTGTAAATTTTCACTGTTTTATTTCTGTTAAAAA
AAA
30 297
MNVHLPLFLASVTLPSICSHFNPLSLEELGSNTGIQVFNQIVKSRPHDNIVISPHGIASVLGMLQLGADGRKKQLAMVMRYG
VNGVGKILKKINKAIVSKKNKDIVTVANAVFVKNAIEVFPVTRNKDVFQCEVRNVNFEDPASACDSINAVWKNETRDMIDNL
LSPDLIDGLVTRLVLVNAVYFKLWKSRLFQNPENTKKRTFVAADGKSYQVPMLAQLSVFRGGSSTAPNDLWYNFIELPYHGESIS
MLIALPTESSSTPLSAIIPHISTKTIDSWMSIMVPRVQVILPKFTAVAQTDLKEPLKVLGITDMFDSKANFAKITRSENHVS
35 HILQKAKIEVSEDGTKASAATTAILIARSSPPWFIVDRPFLFFIRHNPTGAVLFMGQINKP
298
ATGAACTGGCATCTCCCCCTCTTCTCTTGGCCTCTGTGACGCTGCCTTCCATCTGCTCCCACTTCAATCCTCTGTCTCTCGAG
GAACTAGGCTCCAACACGGGGATCCAGGTTTTCAATCAGATTGTGAAGTCGAGGCCTCATGACAACATCGTGATCTCTCCCAT
GGGATTGCGTCGGTCTGGGGATGCTTCAGCTGGGGGCGGACGGCAGGACCAAGAAGCAGCTCGCCATGGTGATGAGATACGGC
40 GTAAATGGAGTTGGTAAAAATTTAAAGAAGATCAACAAGGCCATCGTCTCCAAGAAGAATAAAGACATTGTGACAGTGGCTAAC
GCCGTGTTGTTAAGAATGCCCTCTGAAATTTGAAGTGCCTTTTGTTTACAAGGAACAAAGATGTGTTCCAGTGTGAGGTCCGGAAT
GTGAACTTTGAGGATCCAGCCTCTGCCTGTGATTCCATCAATGCATGGGTTAAAAACGAAACCAGGATATGATTGACAATCTG
CTGTCCCAGATCTTATTGATGGTGTGCTCACCAGACTGGTCTCGTCAACGCAGTGTATTTCAAGGGTCTGTGGAATACAGG
TTCCAACCCGAGAACAAGAAACGCACCTTCTGTCGAGCCGACGGGAAATCCTATCAAGTGCCAATGCTGGCCAGCTCTCC
45 GTGTTCCGGTGTGGTTCGACAGAGTGCCTCCCAATGATTTATGGTACAACCTTCAATTGAACTGCCCTACCACGGGGAAGCATCAGC
ATGCTGATTGCACTCCGACTGAGAGCTCCACTCCGCTGTCTGCCATCATCCACACATCAGCACCAAGACCATAGACAGCTGG
ATGAGCATCATGGTCCCCAAGAGGGTGCAGGTGATCTCTGCCAAGTTTACAGCTGTAGCACAAACAGATTTGAAGGAGCCGCTG
AAAGTTCTTGGCATTACTGACATGTTTGATTTCATCAAAGGCAAAATTTGCAAAAAATAACAAGGTGAGAAAACCTCCATGTTTCT
CATATCTTGCAAAAAGCAAAATTTGAAGTCAGTGAAGATGGAACCAAGCTTACAGCAGCAACAACCTGCAATTTCTATTGCAAGA

TCATCGCCTCCCTGGTTTATAGTAGACAGACCTTTTCTGTTTTCATCCGACATAATCCTACAGGTGCTGTGTTATTCATGGGG
CAGATAAACAAACCC

299

MSQRPRAPRSALWLLAPPLLRWAPPLTLVLHSDLFQALLDILDYIEASLSSESQKYRYQDEDTPPLEHSPAHLNPQANSPPVIVN
5 TDTLEAPGYELQVNGTEGEMEYEEITLERGNSGLGFSIAGGTDNPHIGDDPSIFITKIIIPGGAAAQDGRLRVNDLSILFVNEVDV
REVTHSAAVEALKEAGSIVRLVYMRKPPAEKVMEIKLIKPGKGLGFSIAGGVGNQHIPGDNSIYVTKIIIEGGAHKDGRLOIG
DKILAVNSVGLQEDVMHEDAVAALKNTYDVVYLKVAKPSNAYLSDSYAPPDITTSYSQHLNEISHSSYLGTDYPTAMTPTSPRR
YSPVAKDLLGEEDI PREPRRIVIHRSSTGLGFNIVGGEDGEFISFILAGGPADLSGELRKGDQILSVNGVDLRNASHEQAAI
ALKNAGQTVTIIAQYKPEEYSRFEAKIHDRLQMLNSSLGSGTASLRNPKRGFYIRALFDYDKTKDCGFLSQALSFRFGDVLH
10 VIDASDEEWQARRVHSDSETDDIGFIPSKRRVERREWSRLKAKDWSSSGSQGREDSVLSYETVTQMEVHYARPIILGPTKD
RANDDLLSEFPDKFGSCVPHTTRPKREYEIDGRDYHFVSSREKMEKDIQAHKFI EAGQYNHLYGTSVQSVREVAEQGKHCILD
VSANAVRRLQAAHLPIAIFIRPRSLNVLEINKRI TEEQARKAFDRATKLEQEFTECFSAIVEGDSFEEIYHKVKRVIEDLSG
PYIWPAPERL

15 300

GGATCCGCGGGACAGATGAGGAAGGGGCTTAAGTCACTGCAGCCAGAGGGATGGAGGTGGACTGATGGGAGGGCTTCTCCGGTG
GGGTTAGAAGGAAAAGTAGGAAAGAGAAGTGTAAAGTAGATGGCAGAGGCAGAGACATGGAAGACAGACTCTAGGGTTCCCT
GATGATATCTATCTCGCCCAACACAAAAGGGAGGGTACAGTGGTGGGGGCACCCAAGCTAGGGTGTGAGTACCCTAAGTGTATT
CTTCTGAGATGTAGGCCATTCACTAATCTTGAACAGCTACAGTTTCACAGTAGGAAGACCCCCCAGATTCACTGCCCCCTCC
20 CTTAGTAAAGCCTCTGAGACCTTCTCTGAACATTTCCCTTCTGTCTTTGCCCTCTGTTCCTTCCAGAGACTATGTGCCAGGCAGA
TGGATTCTCTCCGGGCTGAGAGGAAGTGCAGGAATTTCTCTGCCTCTTACCCGTAACCAACCACTTCTCTAGCCCTAGGGCA
GGAAGTCCCAACAAATTTCTACCCCTTTTCTGCAATTTCTATTGGGGTGAGAGGAGGCCAGGAGGAGAGAGAGCTGGGCTCA
GCTTCTTTTGTAGCTGCTGGAGCCCTCTGTGAGGAGGCCCTCTTTGTCTGGCTTCTCAGGAGAGTGTGGCTAGGTTCTGCTGCC
TATGGGAAGAGGGGGCCAGGGTGTGTGGAGCAAGATGGTGGGTGCTGGTGGCTTGGGACCTGGGGGAATGGGACAGCTGGTGC
25 GCTCAGAGACGGCCTACTTTACTCACAGCTGGAATTTAGTGGGAGAGAGCAGCTCAACTCCAATCTGGAGGATTAGGGAGATT
AAAGTGAGAGAAGAGAGATGTCCCAGAGACCAAGAGCTCCAGGTGAGCCCTCTGGCTCCTGGCACCCCCACTGCTGGGGTG
GGCACCCCACTCCTCACAGTGTGATAGCGACCTCTTCCAGGCCCTTGTGGACATCCTGGACTATTATGAGGCTTCCCTCTC
AGAGAGTCAGAAATACCGCTACCAAGATGAAGACACGCCCTCTGGAGCACAGCCGGGCCACCTCCCAACAGGCCAATTC
TCCCCAGTGATTGTCAACACAGATACCTAGAAAGCCCCAGGATATGAGTTGACAGGTGAACGGGACCGAGGGGAGATGGAATA
30 CGAGGAAATCACATTGGAAGGGGTAATCAGTCTGGGCTTCAGCATCGCAGGTGGCACTGACAACCCACACATCGGTGACGA
CCCATCCATTTTCATCACCAAGATCATTTCTGCTGGGCTGCGGCCAGGATGGCCGCTCAGGGTCAACGACAGCATCTGTT
TGTAATGAAGTGACGTGCGCGAGGTGACCCACTAGCGCGGTGGAAGCCCTCAAAGAGGCAGGCTCCATCGTTCCGCTCTA
TGTCATGCGCCGAAGCCCCGGCTGAGAAGGTGATGAGATCAAGCTCATCAAGGGGCTTAAAGGTCTTGGCTTCAGCATCGC
AGGGGGCTAGGGAACAGCACATCCCAGGAGATAATAGCATCTATGTAACAAAGATCATCGAAGGGGGTGTGCCCCAAGGA
35 TGGGAGGTTGCAGATTGGAGACAAGATCCTGGCGGTCAACAGTGTGGGGCTAGAGGACGTGATGATGAAGATGCTGTGGCAGC
CCTGAAGAACAGTATGATGTTGTCTACCTAAAGGTGGCCAAGCCCAGCAATGCCCTACCTGAGTGACAGCTATGCTCCCCAGA
CATCACAACCTCTTATTTCCAGCACCTGGACAATGAGATCAGTCACAGCAGCTACCTGGGCACCGACTACCCACAGCCATGAC
CCCCACTTCCCTCGGCGCTACTCTCCAGTGGCCAAGGACCTGCTCGGGGAGGAAGACATTTCCCGGAGAACCAGGCGAATGTT
GATCCACCGGGGCTCCACGGGCTGGGCTTCAACATCGTGGGTGGCGAGGACGGTGAAGGCATCTTCATCTCCTTTATCTGGC
40 CGGGGGCCCTGAGACCTCAGTGGGGAGCTGCGGAAGGGGACAGATCCTGTGCGTCAACGGTGTGGACCTCCGAAATGCCAG
CCATGAGCAGGCTGCCATTGCCCTGAAGAATGCGGGTCAGACGGTCACGATCATCGCTCAGTATAAACAGAGAGTACAGCCG
ATTCGAGGCCAAGATCCACGACCTTCGGGAACAGCTCATGAACAGCAGCCTGGGCTCAGGGACTGCGTCTTGGCGAGCAACCC
CAAAAGGGGTTTCTACATCAGGGCCCTGTTTGTATTACGACAAGACCAAGGACTGCGGCTTCTGAGCCAGGGCCTGAGCTTCCG
CTTTGGGGATGTGCTGATGCTCATCGATGCTAGTGTGAGGAGTGGTGGCAGGCACGGCGGTCCACTCTGACAGTGAGACCGA
45 CGACATTTGGGTTTCATCCCGACAAACGGCGGTTGAGCGACGAGAGTGGTCAAGGTTAAAGGCCAAGGACTGGGGCTCCAGCTC
TGGATCGCAGGGTCGAGAAGACTCGGTTCTGAGCTACGAGACAGTGACGACAGTGAAGTGCATATGCTCGCCCCATCATAT
CCTTGGGCCCACCAAGGACCGCGCCAACGATGATCTTCTCTCCGAGTTCCCGACAAGTTTGGATCTGTGTTCCCATACGAC
ACGGCCCAAGCGGGAGTATGAGATAGATGGCCGGGATTACCACTTTGTGTCTGTCGGGAGAAAATGGAGAAGGACATTCAGGC
GCACAAGTTTATTTAGGCGGGCAGTACAACAGCCACCTCTATGGGACAGCGTCCAGTCCGTGCGAGAGGTGGCAGAGCAGGG
50 GAAGCACTGCATCTCGATGTCTCGGCAATGCCGTGCGGGGCTGCAGGCGGGCCACCTGCACCCCATCGCCATCTTCATCCG

- CCCCGCTCCCTGGAGAATGTGCTAGAGATTAAACAAGCGGATCACAGAGGAGCAAGCCCGCAAAGCCTTCGACAGAGCCACCAA
GCTGGAGCAGGAGTTCAAGAGTGTCTCAGCCATCGTGGAGGGTGACAGCTTTGAGGAGATCTACCACAAGGTGAAGCGTGT
CATCGAGGACCTCTCAGGCCCTTACATCTGGGTTCAGCCCGAGAGAGACTCTGATTCCTGCCCTGGCTTGGCCTGGACTCGCC
CTGCCCTCCATCACCTGGGCCCTTGGTCTGGACTGAATTGCCCAAGCCCTTGGCTCCCCCGGCCCTCCCTCCACCCCTTCTTAT
5 TTATTTCTTTCTAACTGGATCCAGCCTGTGGAGGGGGGACACTCCTCTGCATGTATCCCCGACCCCAAGCTGGGCTCCTG
AACGCCAGGAACCTGGGGTCTGGGGGGGAGCTGGGCTCCTTGTTCGAGCCCTTGCTCCTTAGGATCCCCGCCCCACCTGCC
CCATGCACACACAGACCCACCGGGGGCCACCTGCCCTCCCCATCCTCTCCACACACATTCCAGAAGTCAGGGCCCCCTCGA
GGAGCACCCGCTGCAGGATGCAGGGCCACAGGCCCTCCGCTCTCTCCTAAGGCAGGGTCTGGGGTCACCCCTGCCCTCATCGTAA
TTCCCATGTATACCTTGATTTCTCATTTATTTTCCACTTTTCTTCTCAAAGGTGGTTTGGGGGGAGAAGCAGGGGA
10 CTCCGACGGGGCCCTGCCCTCCACATGCCCCACCATTTTCTTTGCGGTTTGCATGAGTGAAGGTCTAAATGTGGCTTT
TTTTTTTTTTTCTCGGAATTTTTTGGGAAAAGGAGGGATGGGCTAGGGAGTGGGAAATGCGGGAGGGAGGGTGGGGCA
GGGGTCGGGGTTCGGGTGTCCGGGAGCCAGGAAGACTGGAAATGCTGCCGCCCTTCTGCAATTTATTTATTTTCTTTTGAG
AGAGTGAAAGGAAGAGACAGATACCTGAAAAAAAAAAAAAAAAAAAAA
- 301
- 15 MDRGEQGLLRDPVPEEGEDVAATISATETLSEEEQEELRRELAKVEEEIQTLSQVLAKEKHLAEIKRKLGINSLQELKQONIA
KGWQDVTATSAYKKTSETLSQAGQKASAAFSSVGSVITKKLEDVKNSTPKSFEEKVENLKSXVGGTKPAGGDFGEVLNSAANA
SATTTTEPLPEKTQESL
- 302
- 20 GAGGAGCTCTGCGCGCGCGCGGGCGGATCCGAGCCGGGACGGGCTGCAGGCGGGGGTGTGTCAGAGGACACGAGGCGGCGGGC
TGGAGACATGGACCGCGCGGAGCAAGGTCTGCTGAGAACAGACCCAGTCCCTGAGGAAGGAGAAGATGTTGCTGCCACGATCAG
TGCCACAGAGACCTCTCGGAAGAGGAGCAGGAAGAGCTAAGAAGAGAAGTTCGAAAGGTAGAAGAAGAAATCCAGACTCTGTC
TCAAGTGTAGCAGCAAAAGAGAAGCATCTAGCAGAGATCAAGCGGAACTTGGAATCAATTTCTTACAGGAACAAAAACAGAA
CATTTGCCAAAGGGTGGCAAGACGTGACAGCAACATCTGCTTACAAGAAGACATCTGAAACCTTATCCCAGGCTGGCAGAGAAGGC
CTCAGCTGCTTTTTCTGCTGTTGGCTCAGTCATACCAAAAAGCTGGAAGATGTAAAAAAGCTCCCAACTTTTAAATCATTTGA
25 AGAAAAGGTGCAAAACTTAAAGTCTAAAGTAGGGGAAACCAAGCCTGCTGGTGGTGATTTTGGAGAAGTCTTGAATTCGGCTGC
AAATGCTAGTGCACACCACCGAGCCTCTTCCAGAAAGACACAGGAGAGCCTGTGAGATTCTTACCTTTGTTCTGCTACCCA
CTGCCAGATGCTGCAAGCGAGGTCCAAGCACATCTTGTCAACATGCATTCGCATGAATTTCTACCATGATGCTTTTATTTAGC
TTTACATATTCCTTTGACCAATAGTTTGTGGGTAAACAAAAATGAAATATCTTACCTCTATTCTTGGGAAACACCTTTTAG
TGTACATTTATGTTCTTTATTTAGGAAACACCATTTATAAAACACTTATAGTAAATGGGGACATTCACATATAATGATCTAAGA
30 AGCTACAGATTGTCATAGTTGTTTTCTGCTTTACAAAATGCTCCAGATCTGGAATGCCAGTTTGACCTTTGTCTTCTATAAT
ATTTCTTTTTTTCCCTCTTTGAATCTCTGTATATTTGATTTCTTAACTAAAATGTTCTCTTAAATATTTGAACTCTGGTAA
TTAAAGTTTGGGTGATTTTTCTTTACCTCCAAGGAAAGAACTACTAGCTACAAAAATATTTTGAATAAGCATTGTTTTGGT
ATAAGGTACATATTTTGGTTGAAGACACCAGACTGAAGTAAACAGCTGTGCATCCAATTTATTATAGTTTGTGAAGTAAACATA
TGTAATCAAATCTTAGGTGACTTGAGAGTGAACCTCCTATATCATTTATTTAGCACCGTTTGTGACAGTAACCATTTAGTGT
35 ATGTGTTATATACCACTTATATCAACTTATTTTTACCAGGTAAAAATTTAATTTCTACAAAATAACATTCTGAATCAAGCA
CACTGTATGTTAGTAGGTTGAACATGAACACTGTCATCAATGTTTCAGTTCAAAGCCTGAAAGTTTAGATCTAGAAGCTGGT
AAAAATGACAATATCAATCACATTAGGGGAACCATTTGTTGCTTCACTTAATCCATTTAGCACTATTTAAATAAGCACACCAA
GTTATATGACTAATATAACTTGAAAAATTTTTTATACTGAGGGGTGGTGATAACTCTTGAGGATGTAATGCATTAATAAAAAATC
AACTCATCATTTTCTACTTGTTTTCAATGTGTTGGAACCTGTAAGTATGATACTGTAGAACCTGTCTCCTACTTTGAAAACCTGAA
40 TGTACGGGCTGAGTGAATCAAAGTGTCTAGACATATTTGCATAGAGGCCAAGGTATTCTATTCTAATAAGTCTTACTCAACAC
TACCACCTTTTCTTATACGTATATGATTATGGCCTACAATGTTGATTTGTTATTTATTAATTTGATTTGTTTATTTATTG
TTTTATGCCAAATGTTAACTGCCAAGCTTGGAGTGACCTAAAGCATTTTTTAAAGCATGGCTAGATTTACTTCAGTATAAATTA
TCTTATGAAAACCAAATTTTAAAGCCACAGGTGTTGATTTGTTATATAAAATACATGCTGCCATTCTTGATTGCTAGAGTTTTG
TTAGTACTTTGGATGCAATTTAAACTATGTGCTATCAGTGTGAAAGCTTAAATAATTCATCTATCAGTAGTATAGGTCTCA
45 ATATTTATTTAGAGACCAGTGGTCTGGAAACAGCTTGTGTTACCGAATCAACTGGAGTCTATGCTTAAAAAAGGGCTTTAAGGCACAGGCTGTTGAAGCATT
TTTCAACATCTTAAATTTATGCTTAAATGGTATCATATTAACATATTTCTAAATAAGGGCTTTAAGGCACAGGCTGTTGAAGCATT
TTCTCAGAGGAGTGGATCTGTAGAAGTCTGTCTTTCTATAGAAATATTGTGCTTACTCAAGTGTAAATTTATTTTCTATGAA
CTAGTCTACTTCTTAAATTTCAAACATATTTCTTTGATCATTGTTTCTTGAGCATCTGCCCCGCTACTAACCTTTTCAACAA
GGCAAAATGGAGTAAAGTGCAATTTCTTTAGATGAGTGAATACCTCAAGTCTCTTTTCTGCCCAAAAGGGGAAAGTGATA
50 GAAATGGGGGTGGCAAGTGGGGTGAAGTGGGATGAAGTGGGTATTGGGGGTGGCTGTGAAGAAAATAATGGAGAATCACTTTTCT

TAGACATCTACCTATACTTAATCTAAGAAACAAAGTAATCTACTGTAAAGTACTCTGCCCTTGAAAGAAGTATTAAGAGAGT
GAGGATGGATTTAGAAAAAACAATGAATTTAGAAATATTCAAAATGGTTTTTGTGGCAGATTCAATATTATGAATTCACAGATA
TTTAAAGAATGAGAAACATAGTAATTAGTAGAAATGCCAGAAACAGTTCCTGGTTCCTCTGTGTTTGACACTAAGAAAAATAGC
AAGAGTGTGAAATCTCAGATACTTATGAAATCTCACAGATGTAAGGACTCAAGTGTAGAAGAAAAATATCCCCTTCTTACAAAA
5 GAAATGTCAATTTATGGAGTTTGTGGGAAATAGGGCAAGAAATCTTATGCTTATGAGAGCCAAGTAGTCAGTGGAGAGAGTAG
AGCTCAAACTGGATTATCACCTTAGCAACTTAGAATAGTTTGAATAGAAAAAAGTATTTAATTTGGATCTGGATCTGTTAA
GATATGCACAGTCTATTTTTTGTATAGTATTGGAAAAATAAAATGCTATAATTTG

303

MSNVVRVNSGSPSLERMDARQAHPKPSACRNLFQVPVDHEELTRDLEKHCARDMEASQRKWNFDQNHKPLEGKYEWQVEVEKGS
10 PEFYRPPRPPKGAACKVPAQESQDVSGSRPAAPLIGAPANSEDTLHVPKTDPSDSQTGLAEQCAGIRKRPATDDSSSTQNKRA
RTEENVSDGSPNAGSVEQTPKKPGLRRRQT

304

ATGTCAAACGTGCGAGTGTCTAACGGGAGCCCTAGCCTGGAGCGGATGGACGCCAGGCAGGCGGAGCACCCCAAGCCCTCGGCC
TGCAGGAACCTCTTCGGCCCGGTGGACCACGAAGAGTTAACCCGGGACTTGGAGAAGCACTGCAGAGACATGGAAGAGGCGAGC
15 CAGCGCAAGTGAATTTTCGATTTTCAGAATCACAAACCCCTAGAGGGCAAGTACGAGTGGCAAGAGGTGGAGAAGGGCAGCTTG
CCCGAGTTCTACTACAGACCCCCCGGCCCCCCAAAGGTGCCTGCAAGGTGCCGGCGCAGGAGAGCCAGGATGTCAGCGGGAGC
CGCCCGCGCGCGCTTTAATTGGGGCTCCGGCTAACTCTGAGGACACGCATTTGCTGGACCCAAAGACTGATCCGTCGGACAGC
CAGACGGGGTTAGCGGAGCAATGCGCAGGAATAAGGAAGCGACCTGCAACCGACGATTTCTTACTCAAACAAAGAGCCAAAC
AGAACAGAGAAATGTTTCAGACGGTCCCCAAATGCCGGTCTGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAAGACGT
20 CAAACGTAA

305

MCFSFIMPAMADILDIWAVDSQIASDGSIPVDFLLPTGIYIQLEVPREATISYIKQMLWKQVHNYPMFNLMDIDSYMFACVN
QTAVYEELEDETRRLCDVRPFLPVLKLVTRSCDPGEKLDISKIGVLIGKGLHEFDSLKDPEVNEFRMRKRFSEEKILSLVGLSW
MDWLKQTYTPEHEPSIPENLEDKLYGGKLIVAVHFENCQDVFSFQVSPNMPNPKVNELAIQKRLTIHGKEDVSPYDYVLQVSG
25 RVEYVFGDHLPIQFYIRNCVMNRALPHFILVECCIKKMYEQEMIAIEAAINRNSNLPLPLPPKKTRIISHVWENNPFQIV
LVKGNKLNTEETVKVHVRAGLFHGTTELLCKTIVSSEVSGKNDHIWNEPLEFDINICDLPRMARLCFAVYAVLDKVKTKKSTKTI
NPSKYQTIKAGKVHYPVAWVNTMVDFKGLRTGDIILHSWSSFPDELEMLNPMGTVQTNPYTENATALHVKFPENKKQPPY
YPPFDKIIIEKAAEIASDSSANVSSRGKKFLPVLKEILDRDPLSQLCENEMDLIWLTRQDCREIFPQSLPKLLLSIKWNKLEDV
AQLQALLQIWPKLPPREALLELDFNYPDQYVREYAVGCLRQMSDELSQYLLQLVQLKYEPFLDCALSRFLERLALGNRRIGQ
30 FLFWHLRSEVHIPAVSVQFGVILEAYCRGSVGHMKVLSQVEALNKLKTLNLSLIKLNKLNRAKGEAMHTCLKQSAYREALS
DLQSPNPNPCVILSELYVEKCKYMSKMKPLWLNVNKNVFGEDSVGVIFKNGDDLQDMLTLQMLRLMDLLWKEAGLDRMLPYG
CLATGDRSLIEVYSTSETIADIQLNSSNVAFAAFNKDALLNLWKEYNSGDDLDRALIEFTLSCAGYCVASYVLGIGDRHSDN
IMVKKTGQLFHDIFGHILGNFKSKPGIKRERVPIILTYDFIHVIQGGKTGNTEKFRFRQCCEDAYLILRRHGNLFIITLALML
TAGLPELTSVKDIQYLKDSLALGKSEEEALKQFKQKFDALRESWTTKVNWMAHTVRKDYRS

35 306

ATGTGCTTCAGTTTCATAATGCCTCCTGCTATGGCAGACATCCTTGACATCTGGGCGGTGGATTACAGATAGCATCTGATGGC
TCCATACCTGTGGATTTCCCTTTTGCCCACTGGGATTTATATCCAGTTGGAGGTACCTCGGGAAGCTACCATTTCCTATATTAAG
CAGATGTTATGGAAGCAAGTTCACAATTACCCAATGTTCAACCTCCTTATGGATATTGACTCCTATATGTTTGCATGTGTGAAT
CAGACTGCTGATATATGAGGAGCTTGAAGATGAAACAGAACACTGTGTATGTCAGACCTTTTCTTCCAGTTCTCAAAATTAGTG
40 ACAAGAAGTTGTGACCAGGGGAAAAATTAGACTCAAAATTGGAGTCTTATAGGAAAAGGTCTGCATGAATTTGATTCCTTG
AAGGATCCTGAAGTAAATGAATTTTCAAGAAAAATGCGCAAAATTCAGCGAGGAAAAAATCCTGTCACTGTGGGATGTCTTGG
ATGGACTGGCTAAAAACAAACATATCCACCAGAGCATGAACCATCATCCTGAAACCTTAGAAGATAAACTTTATGGGGGAAAG
CTCATCGTAGCTGTTTCAATTTTGAACCTGCCAGGACGTGTTTAGCTTTCAAGTGTCTCCTAATATGAATCCTATCAAAGTAAAT
GAATTTGGCAATCCAAAAACGTTTGACTATTATGGAAGGAAGATGAAGTTAGCCCTATGATTATGTGTTGCAAGTCAGCGGG
45 AGAGTAGAATATGTTTGGTGATCATCCACTAATTCAGTTCAGTATATCCGGAAGTGTGTGATGAACAGAGCCCTGCCCCAT
TTTATACTTGTGGAATGCTGCAAGATCAAGAAATGTATGAACAAGAAATGATTGCCATAGAGGCTGCCATAAATCGAAATTCA
TCTAATCTTCTCTTCCATTACCACCAAAGAAAAACGAATTAATTTCTCATGTTTGGGAAAAATAACAACCCCTTTCCAAATTGTC
TTGTTTAAAGGAAATAAACTTAACACAGAGGAACTGTAAAGTTCATGTCAGGCTGTGCTTTTTCATGGTACTGAGCTCCTG

30 MRAVLDTADIAIVALYFILVMCIGFFAMWKSNRSTVSGYFLAGRSMTWVTIGASLFVSNIGSEHFILAGSGAASGFVAGWEF
NALLLLQLLGWVFIPIYIRSGVYTMPEYLSKRFGGHRIQVYFAALSLILYIFTKLSVDLYSGALFIQESLGNLYVSVILLIGM
TALLTVTGGLVAVIYTDTLQALLMIIGALTLMIIISIMEIGGFEEVKRRYMLASPDVTSILLTYNLSNTNSCNVSPKKEALKMLR
NPTDEDVPWPGFILGQTPASVWYWCADQVIVQVRVLAAKNIAHAKGSTLMAGFLKLLPMFIIVVPGMISRILFTDDIACINPEHC
MLVCGSRAGCSNIAYPRLVMKLPVGLRGLMMAVMIAALMSDLSIFNSASTIFTLDVYKLIRKSASSRELMIVGRIFVAFMVV
ISIAWVPIIVEMQGGQMYLIQEVADYLTTPVAALFLLAIFWKRCNEQGAFYGGMAGFVLGAVRLILAFAYRAPECDDQPDNRPG
FIKDIHYMYVATGLFWVTGLITIVISLLTPPTKEQIRTTTFWSKKNLVKENCSPKEEYQMQEKSILRCSENNETINHIIPN
35 GKSEDSIKGLQPEDVNLLVTCREEGNPVASLGHSEAETPVDAYSNGQAALMGEKERKKETDDGGRYWKFIDWFCGFKSKLSKR
SLRDLMEEEAVCLQMLEETROVKVILNIGLFAVCSLGI MFVYFSL

CGGGCCGCCCGCGGGCGGAGAGGGGGCGGCGCGGCGGCAAGCGCGCGGGCGGGGCGGGGCGCGCCGAGGGGGCGTGCGCGGGC
GTCTGCGCAGCTGCCAGCGCCTTTAAGCCCGGGCTCGCGCTCTCGGACCGTGCTTTTCGCCGCGCTGGGAGCCGTCGCGCGCAGCA
40 GTTTCTAGGTCCCCACTGTCCCCGCCGTCCGCCCCCTTCGCGTCCCGGGAACCGGTGGCTTCGAGCCGCACTCGCCGATCTT
CCAGGCATGCCCCGCTACGAGCTGGCTTTAATCTGAAAGCCATGCAGCGGGTAAGTGACCTTCCTTCAGAGCCGGTTTTCCCG
CGCGGGCGCCCCCTGTGCTTGTCTATGACCCCTATATTAGGTTGGCAAATAGATGTATGTTTATGTGTTGAATAAAATAATTAA
TAAATAATAATAAGGTACGTATATATAGTPTTATGACTGGATATGCTTTTGGGTATGAAGGCTTCTGCTTTATAGAAAAATGCTG
ATTTTTTTGTTTTCTATTATAAATGAGTAGAAAAACCTTGTAAAGATCTGAAATGAAGATACTGGATGTGTTACCATAGAATAAAA
45 CAAAGACTTGGCATTGGGAGACTAGAAGGAGCTTTAGAGGCAGATGGGACTGGGTTTTTGGATTGTAGCACCTACTTCTTAAGGAC
TTTAACCTTCTGGATCTTCAGTTCCTTCCTTTGTAAAGCCGGAAGTAATTCATTGGGAGGGTATTTGAAGATAAAGTGGTTATGT
GACCTTGGAGGTAGTTAAGTGATGTTCTGTCTATTAAGTAGCTTTGTGGTCTTAGGCAAGTCGCTGTGAACCTCATTGCCTAGT
TTTTCTATPTATCCCACTTTATTAATGAGGAAGAACTCTAAAGATAAAACTTATTATATCAAGTTTATTTCAATCAGAAATCCCAA
CTAGATCAAGCTGAAATACAGTTGCTTGTATGTAAGTAAATGTTGAAGTATGCGTGTGGTGCTTACTACTCTGCTCCCATCTAT
50 TATTGCCATGAGGGGTGTGGGCCCTGGTTTTGCCAGCTTTTTTTTTCTCTCTTCCCTCTACAGCCAGAAATCCAGATTTTTATAG

GAAATCTGGTTTTTTTTTTTGTATGTTGCCAACTAATTGAATACTTAGGTACCATGGATCAAATGTAATGTATCTGTGGGTTAG
GTTGGCTCTCATTTTGTAACTTAGATATTCAGATAATTAGCTGTAGGTTGTCTGTAAATCTCTTAGCATTTATGTGACAAAGC
CCACCTTGTTAATTGGTCACTGGTTTGCAGTAGGACTGATTTTCAGCAATGGCAGGCATATTCAAATCTTGAAAGTATCTATAGT
TTCCGATTTGTTGGGTTGTATATGTGCATTTAGGCTAAAAATATACATGTAGGCTTCAGAGCTGCACCGTGTAAACTGTGAGGGAG
5 TGACATTCACAGTCCACTTTTTCTTTCAATGTCTGAGTGATATCTGTTTTAGGCATGTATTAGTACTATTTGGTTTGTACCCA
TTGTATTAACTCTTAGGTGTTTGAATACTACTGGCATGCAAAGATTGCTTTGCCTAGAAAAATTTTTTTTTTGTCTCAGAGTG
TATCTGTGGAATTAATAAGTATTACCCATGATCCCTGGAAAAATAGTATTACCATGAAAAAATTCAGGTGCTCTCTTGGTAAA
GATTATAGAGGAGGAGGTGATGACATTTTATTTTTTGACAGGCTTTGTGTTTCCATAAAATAGAAGCCAAGTTATTACCTTG
ACTAACCTTCCTTTTTTGTCTTTTTTTTCACTCCCACTCCAGATAGATAACATCAAATTGCAAATAAAACGAACCTCAGATCCCA
10 GGTGTGTTCTTACGTACGTTTTTGGCCCAAGGTGPAATTGACTTGTTTTTGCTTCCCTTGCAATCATTTTTACTATGATTACAGTC
CCAGGTAAGGTAATTTCTGTCCACTTTGGGCATACCTCAGTAGAGGTTGGTGTAGATGAAAGCTTAAAAATGAAATTGCAAGTG
AAAGCACAGACTTAGCAAAAACCTCTTTATTTGGAGCCCTCAATCTCATAAACTGGCATTTTTCTGCTTCTGCCCGAGAATCGAG
AATTCCTGGTTTTGTCTTTGTTTTTGACTTGAGTTTTTAAAGTTTCTCTGTAACACCTACCCCTGGAGTGGTTAAGTGTCTTAAT
TTTTGTCAAGTGAAAGAAACCAAGTGCAGTTATGGTAGCCGGTTGCTGTGGATATAATCCAGTTTGAGAAAAATCTCATTTGTGGT
15 GATCAGAAGTGCAGTGTGTGGTTGCATTTTCAGTGTGCGGAAGATGCTCTTGCAAATCCATGTGTCTGCTTTTGGCTTCCGTTAG
AAAGGTGCAGAATCAGTGTAGTTCCAGATATCTGCTCCCTTTTCATTTTTGTGCTCAGATGCCCCGTGTATGAATTCGAATAT
AAAACATCTGTTTCAGTCTGGTTTTGATTCAGATTTAAGGTGAATTTCTGTTGCTCAGTTGAAAAGCTAAATTTGAGTGTAGGCT
CTGTCAGGCATGTTGTAGGAAAGGTGTATGGTACTGAAGTGTGTGGTTTTGTTTTTAACTGCTTGACACTGTTAAGTAAGCAA
ATGACCCAGTCATGGGTTGTTTTGATTGCCATTTCTCTGTTGGTGTACTGCTTGCCCTTCTGTTTTGTTAATAGTCATATAGAGAA
20 AACAGTTGTATGAAGTATAGTATCTTGAAGCTTTAAGACACCAAGTTTCCATTAGGCAGCCTAAGTCTGCGAGACTTTGGTAGC
ACGTTGGTCTCATGGGAATGCCATTCTCATGTGTGGACTAAGTAAGGGCCAGGAGATTTTGCAAAATTTGTGTGTGTGAAGAG
ATCGAAGAAAACATTTTTCTTAATGTCACTGTCTTGAGTATTTTATCTGATTAGGATGGACTCATTTGAAAACATTATGATAATTT
CAAGATTTCTTGTCATCTGGCATACAAAGAAAACAGTCGTGACCCTATTTCCCTGGATTTTTCTATCCCTCTACTTGTCTGTG
TTTGAATGCAAGAGGAAAATCCAGCTGCCTTTATAATCTAGAATACCATGTTTCTGTTTTTATTTCTTAGGTTTAATTCAGGG
25 CGTAATTTCTACAGTCCCTCCAAATACGTGTTAAAGTTACTCTTTTTGTGGTCTTACTTCTTACTTCTTGAGTGTATATTCTC
CAAACCTTGGAGAAAATAGAAGGTTTTAAAGTTTTCCAACCTGAAAAGTATCTTTCAGAGGAAATACCTTAATTTCTGAATCTTT
CTATTTATAACTTGGCCTGGAGAAATTATAATCTGGTCTACCAGGAAGGTGCTTGTGCTTAGGTTTTCTGCTTAGTCACATGGA
AAGATGAGCAGAAAGGATTCCTTTCTAGAAACAGTTTTTTTCTAGAAACATAAAATTGGAACCAAGTAGACTGAGAACAGAAAAA
AAACACAAAGTTGACCTGGTAGGATGTTTTACCTCTATAAACCTCATAGTTGCTAGCCTTTGAGGGAGGCTAGATGAGTTGATA
30 ATAATGCCTAACCTTAGCAATGGATTGAAGTAGGGGTGCAGGGTTATGTAGGTGTGTATATGTATTTAGGATATATATACACA
CCTACATATTATACACACACACATATGTACAATTTCTCTTGAATCCTCAGCCATCTAAGGTTAGGTAGAGCTATTTCCGTTGTTGG
GAATGGA AAACTGAAGAATTAAAAACACAGACTTGGTCTACGTAATGGTAAATACAGTAGTTAATGAAATTTGTATTGTAC
CTATCACTCAGCTTCAACTGTTAATTTCCAGTCTGCTCACCTCTGTTTTCTTTATCTTAAAAATTTTTTTCATTAATTTCCCTTG
GCTAACATGACTAAAAAAATTTACAGCTTAAAGAGCAGTTTCAATTAACACTTGTATGATTTTTGCTTAGATCACTAATTTGTAG
35 AATATGCCTTATTTTATCTTTTTTAAATCTGCGTTATTTTCAGAGAGTGCAGACATCTTGACATACCTTTCTTAGGAATAAGGGC
CTTCTCTCAACCAACATAGTCATCACATTCAAACGGACATTGATACGATACTACCACCAATATTCATTCATATAGAAA
TTTACGCTGTTCTCCAGTAATAGTTTTTAGTGTTTTAATTCAGTATTCAATCAAGGATCATATATTTACTTGTTTAGTCTCCT
TGAACAGTTCCAAAGCCTTTGTCAATTAACAGTGGCACTTTGAGGAGTTTTGGAGTCCAGGCCTACTGTTTTGCAGAAACCC
TCATTTGGAGTTGACTGTAGTTGCGTTTTTCATGATTAGAGTTAGATTAGCCATTGTTGACAGGAATGAGTATTTTCATAGGGAT
40 GTCAATGCATCACACCGGGAGGCATACCACTTTGTTAAGCAGTCATTTGGTTAGGTGGCATTACCCAGAATGGGCCATTGTAAA
TGTACCTATTTCCCTTTTGTATTAAAGTATCTATGGAATAACAACTTGAGACTGAATATTTGTTTTTCAACAACTTCGATTT
GGTTTTAGCATCTGTTGGTGAAGTGAATCAGTTATTCAGGAGTACTTTAAAGCAAATTCAGATATATCATTTACCTGTAAAGTA
ATTAGTGATAAGAAAATAAGGACTTTTTTTTTTTTTTTTTTGTAGACTGAGTCTTGCTCTTTTCGCCAAGCTGGAGTGCCAGTGG
CGTGATCTCGGCTCACTGCAAGCTCTGCCACCGGGTTTCACGCCATTTCTCTGCTCAGCCCCAGCTGGGACTGGGACTACAG
45 GCGCCCCCACCACCCAGCTAATTTTTGTAGTTTTAGTAGAGACGGGGTTTCAACCGTGTAGCTAGGGTGTGCTTTGATGTC
CTGACCTCGTATTCACCCGCTCGGCCCTCCCAAAGTGTGGGATTACAGGTGTGAGCCACTGCGCCCGCAGAAATAAGGAC
TTTTAAACAAACCGTAGTGCTTTTACACCTAACTACAATTCCTTAAACACCTCATGTTTCATGTTTCAAAATTCCTGAAATTTGCAAAA
GTGCTGTCTACAGTTTTGGTTGACTCTGAATCTAAACAAGGTCCACAAATATGTTTTGCTTAATGACCTTAAGTCTCTTACTC
TGTAATAGCTTCTTACCCCACTCTTTACCTTTCCGATTTGGCTGGTCACTTATGCTGTTATATATATTTCTTATTCAGGT
50 AATTTCTGTAGACCATAGGTTCTAGAAAATTTGATTAGATTACAGTTTTGTTTTGTTTTGTTTTGTTAAGAGCACTTAAGTGGTC
TTGGGAGGGATAGATCACATCAAGAGGGACATGTCTGCATTTTAGAGAAAAATTAATCTGGTGTGGGTACAGATGTTAGCA
GCTGCATGCGTCCATTTTGAATCCCCCAGCAGCTTTCTCTGAAAATTTGTAGCATCCCTGTACTCATAGCCTGGGTGCATTAC

GAGATGAAGTTGTTGCCTGAGATCCTACAGTTAGGAAGTGACGGAGCATGGAACCCAGCCCTTAAAACTGTGCTGTTTCTTTC
TGTACCGTGCCACCTACATAGATGTGGTACCTACACCCACCTCCCTTCTTACAGAATCATGAAATCTGTCTGTTGGGGTGGGGA
GAAAACAAGCTTGGGTATGCCCATGTTTTTTGTTTTTGGTAGGTTACACAAAAGTGATTGTAATTCAGGTAATTAGTTTTTCAGA
GGTGTCTGTGTGTTCTATAACCCCTCAATTATTAGCTGGCTAAGTGAATATTTTTCATTTGGATCAGAGGATGAGCAGGGCTCA
5 TCAAGCTTACCATTGACAGAGTGCTGGACTAGGCGTCTCAGATTGTACCTTGAGTGTGTGGTTTGTAAATTGACTCCAAAGCTA
AAAGAGGTGGTTGTTTGGTCACGTTTAAGATTATTGCGTACTACTCTTTGTCAGGCCCGCCCTGTGTCAATTGAACTTATGACTT
GAAGCCCTGAGTATAGGAGATAATTTTTCTTTGTTCCCTTGTTTGGTTTTAGTGGCAGGTGGTTTTTTATTTTATTTATTTTGA
TAGAGTGCCCATATATCCTGCTGTTTACATATAGTTCCAGAATAAGTTTCTGGTTTGGAAAGATAAAATTATATGGTCACCCAGAC
AGTACTAGAGAGATCTGGTCATGATAACTAATTACATTTTGGTTAATCATGGTTATATCACCTGAAGGAGTGCTTATTTAAATTT
10 GGTTCTTACCTTTTTTCTCTTTCTATTTGGTCAAATGCAGCAAAATTTTATATTTACACTTCTCTATTTGGATTAAATCTTTCAG
TTGCAATTTAATTTATCCAGATGCGTGGAATAGAACCAAACTGCATTCTAATTTCTGTGCAAAAACAATAAAATACTACTTTCTG
AGAAGGATGCCCTTCTCTTTTGGGAAGAAGGGGAATCTGCTCTGTTTCAAAGCACCTTTGTGTTCAAGGTTTGGCTAAATCTGCT
CATCACCACTAGGGGACTGTTTCTTTTGCATTCTACAAGTATTACTCCCATACAGAAGCTTTTTAAGAAAAAATGGACTAC
CAAGGAACCCACCTACATGCTTTTCAATACTATACCTATTTTATGTCTTATAATTTGACTTGAGAGGTTTGGCTGAATGATGA
15 ACTGTGACAAAAAGCATATATAAATATAAAATACAAGTTAAAAAGCCCTCATGTTAGCTGAAGATTTTCCATGTGTATTCTTAT
AAAAATAAGGAACCTGGACAGTGCTCATTTGTATTTTGTGAAGAAATGTTAACTTCTTTAAAGGGAGGAAAAACATATATAT
ATATATATAACATCTGTATATTGAGAACCATTTTAAAGGATGTAGTTTATCTCTTGAAGACTGAAATTCAGGCTGTGACT
ATGACTTAGAGTGTAATAATATAGCTGCATTGCAATTATAGTATTTTATATATTTGTACAAAACAAATTGACCTGTTTTTGGTTTTG
TTTTTCCCTTTTCCAGTCTTTTACCAGGTGCTGGTCTCTGCATGTGTTTTGTTGGAAGTTGGAATAAGTAGACACGACCCACAGAG
20 TGGTATGGTATCCAGTGAGAAGGGAAGTGTCCGAGTGCGTTTGAAGGATATGAGGAAGCCTAGGCGGAGAGTTGAGAATTCAG
CCTTGAGCTAGTGTAGGGTACTATTTGGTACCTCTCCTGTGGGCAGAGCGACTTTACTAGAGAAGTTAGGAGTTAGGGCTAGC
ATGAGAATTCATGCAGTTTGGAGCATTCATCAGGCAGCAGAATATAGCTTCTATATTTACCTTGCATAAATCAGATTTGACAAAT
GAAGAACAATGTTACTGATAGTAACAGTGCTTCCAGTGATTGAGGAAGTGCTTGTTTTCTTTTTCAGCACTTTGTCTATGAA
CAAATAAGGTGCAGCAGCTAGTATGACACATAAACCACCGTATTTCCCCCAACGGCCAGATGAAACCAGAAATACCCACTGT
25 GTTTTTCTCTCCCTTCTTTTGGAAAGCAGTGTTGGTACTGCCATTGGGAAGAGTGATACCTGATTAGCCAGGTAGCCTTCTGA
AAATTGAAACAAGTGCCCTGAGGATGACTTGGATGGCTGTGGGTAGCAGAGAGCCCTGGAGGCAGGGTTTCATAACTGCGAAC
AGCAGTTTTTCATTGTAATTGACACAGCAGATTTTCATTGTAATTGACAAGCTTCCCTTATGCAGAAACATTTAACGATACCTTTT
TAAATAATGTTTACTAGTCATCTGATTAGATTTTATTAAAAATTTAGATGATTCGCTCTACATAATGTCAAAGCATTTGTTTAA
AACTTGGAAAGCTTTTGGGGGTGGGGAAGGTGGTATAGAAATATGATCATCTTCAGGATTTGGCAGGTATGAAGTGATACAAA
30 CCTCAACTTTTAAAGATTTTTTTTTTTTTTTTTTGTCTTCTCTGGCAATCAATCTTATTTATGCCAGCCGTTGACACTTAGAAAG
GATTTAAATCACTTTTTCATAGCATAAATTATTGTATTGCTTAATGCTTTAAAAAGGCACGCAGATGGGATTGAGACATTTCTATC
CAGGCAGTTTTTCTGCAATTAGACCTAAGAACCATAATTTGTAGGGCAGTAGTCTCTAGGCATTATGTTTAGCTCCTTTTGTCA
GTGAGAAGTGTTGTAAAGTTTCTAATACTGTGGTCTGTTTTGAACGGTTTTGAGAAGAGGTAGCTTTAGTATAATTTATAGG
ACTTAAAGGAAAAAGCCTAGAAATCATCAGAAAAAGCATGGAATAAATCTTCTCTGAGGTTGCCCTATTGGAAGGAGGAAAA
35 TCAATTAGAACTTAATGTGTGAAGGCATTAGGTTGAATAAATTTGTTGCGTATTGATCTGTCAGCTGCAATTAACAATTAAC
TGAAAAATTTAGTATTTGTCTCTTTTTTCAGGAGAGAAGCCGTCAAATTTTAAATACATTTAATTTTAGGTATATTTAAATGTCTGT
TTTGCAGAGTGATTATCTTAATCCAGAACTGAAAAATTCCTGATGGTCTTACTGAAGATTTATCTAATGGTCCACAAATGTCTG
TTTTGTTGAGGTTTGGCGGACATAATAATAAATCATTTTTTAGTCTTCTGAAAACTTTTCTCCTCCTCCCTCTTTTCTCTGCTT
CTTATACCTTCAACAAAAAGCATTTAAAGGAAAAACATAAAATCACGCCTTTAGTCAGTGCCGTTATCCTCTCAGAGGTGTTTT
40 TGGAAATACTGCCTGATGAAAAGGAAGAGAAGGAAATGCTTCTTTTAAAGTTTGTCTTTTATCTTCTTGCAGCATCTTCTAAC
TGGCAGAAATAATACAATTTTTGAGTTGTTTTGTAAGTTGGATATTTTTAGTTTTGTGTGGAATGTCTGTTTCCCATTAGGGC
AGGGGGCGTCTCACAAGATCTAGGTTACTGTATGACAACTATATCCATTTACTGGAGTTTAAAGAGCTTTATAACAATACATTT
CCAGAGACTAGGCAGTACTTGGTTTTCTGTATTTAAATTTTCTGAGGGTAGATGATAACATCTAGTAAACACCACCTGCGCTGT
CCATCCAGGACCCTAACCTGGCCATAAAGTTACCCCCAACTCTCCTAGGATTCATAGTATTTTAAAGATTTTCATTTTTGTGCCCT
45 TATTACGAATATAATTAGTGTTTTCTCCCTTTAGGGTGTCTTAAGATCTTAGTGCTGTCTTAATCTGAGCATACCAAGTAA
GACAAGAAAGGGCAATGCTGTGGTAGGCAGTTAGCTGTGAAGTTAGGCCACTTAGTTATGAGGTAGATGAATTTCTGTCTCA
GCTCAAAATGATAATTTCTCACTGACAAAAGTATCAGCTGTGATCATATAGATTGCTCTTCTCTGTGGAAGTGAAGTGATTAA
AAAAAATTTTTTTCTTCCAATAAAGAAATGTCTCGTTGCGCCGGGTTTGTAGACTGAAAGTTTCTCATCACGTGGCGAGACC
TCCCTACTTTCAATCTTTGTCTTTTGTCTCTGCTGCCCTGTCAGGGTTGGTACAGTAGGCTTCACTAGACTTAGCTGCAACTCAG
50 AATTTCTCCTCCAGCACCTGAGTAAATGCTGATGGTCTTGTGGAGAGTGGATTAAAGAGTACGAGCTAAGTTCTCAATCCCAATT
AAGAAGCGGAAAAATTTAACTGTCTTCTTCAAAGTTTATCAACAACCACCACCATCAAGACAGCAAAACCAAGGACAAAGACTTT
GACCCGTGCTGTGTGCTCTGTGTAGTCCAGTTACAGTATGGTTTACAGACTTGGCTGGGGTTACTAAAAATTAATAAAAGTTG

GACACTTCTGTCTATTGGAGCGCTATTATTACAAAGTTACCAGAATGAGAGCTGTACTGGACACAGCAGACATTGCCATAGTGGC
CCTGTATTTTATCCTGGTCATGTGCATTGGTTTTTTTGGCCATGTGGAAATCTAATAGAAGCACCGTGAGTGGATTTCTCTGGC
GGGGCGCTCTATGACCTGGGTAAACAATTGGTGCCTCTCTGTGTGTGAGCAATATTGGGAGTGAGCACTTCATTGGGCTGGCAGG
ATCTGGAGCTGCAAGTGGATTGTCAGTGGGCGCATGGGAATTCAATGCCTTACTGCTTTTACAACCTTCGGGATGGGTTTTCAT
5 CCAATTTACATCCGGTCAGGGGTATATACCATGCCTGAATACTTGTCCAAGCGATTGGTGGCCATAGGATTAGGCTATTTT
TGCAGCCTTGTCTCTGATTTCTATATTTTACCACAGCTCTCGGTGGATCTGTATTTCGGGTGCCTTTTATCCAGGAGTCTTT
GGGTGGGAATCTTTATGTGTCTGTCATCTGCTCATTGGGCATGACTGCTTTGCTGACTGTCACCGGAGGCCCTTGTGTCAGTGT
CTACACAGACACTCTGCAGGCTCTGCTCATGATCATTGGGGCACTTACACTTATGATTTATTAGCATAATGGAGATTGGCGGGTT
TGAGGAAGTTAAGAGAAGGTACATGTTGGCCCTACCCGATGTCACTTCCATCTTATTGACATACAACCTTTCCAACACAAATTC
10 TTGTAATGTCTCCCTAAGAAAGAAGCCCTGAAAATGCTGCGGAATCCAACAGATGAAGATGTTCTTGGCCCTGGATTCACTCT
TGGGCAGACCCAGCTTCAGTATGGTACTGGTGTGCTGACCAAGTCACTGTCAGAGGGTCTTGCAGCCAAAAACATTGCTCA
TGCCAAAGGCTCTACTCTTATGGCTGGCTTCTTAAAGCTCTGCCAATGTTTATCATAGTTGTCCTCAGGAATGATTTCAGGAT
ACTGTTTACTGATGATATAGCTTGCATCAACCCAGAGCACTGTCATGCTGGTGTGTGGAAGCAGAGCTGGTTGCTCCAAATATTGC
TTACCCACGCCCTGGTGATGAAGCTGGTTCTGTGGGCTTTCGGGGTTTAAATGATGGCAGTGATGATTGTCAGCTGTGATGAGTGA
15 CTAGACTCTATCTTTAACAGTGCCAGTACCATATTCACCTCGATGTGTACAACTTATCCGCAAGAGCGCAGCTCCCGGGA
GTTAATGATTGTGGGAGGATATTTGTGGCATTTATGGTGGTGATCAGCATAGCATGGGTGCCAATCATCGTGGAGATGCAAGG
AGGCCAGATGTACCTTTACATTACAGGAGTAGCAGATTACCTGACACCCCGAGTGGCAGCCTTGTTCCTGCTGGCAATTTCTG
GAAGCGCTGCAATGAACAAGGGGCTTTCTATGGTGGAAATGGCTGGCTTTGTTCTTGGAGCAGTCCGTTTGATACTGGCCTTTGC
CTACCGTGCCCCAGAATGTGACCAACCTGATAATAGGCCGGCTTCATCAAAGACATCCATTATATGTATGTGGCCACAGGATT
20 GTTTTGGGTACGGGACTCATTACTGTAATTGTGAGCCTTCTCACACCCTCCCAAAAGGAACAGATTGCAACCACACCTT
TTGGTCTAAGAAAGAACCTGGTGGTGAAGGAGAACTGCTCCCCAAAGAGGAACCATACCAAATGCAAGAAAAGAGCATTCTGAG
ATGCAGTGAGAATAATGAGACCATCAACCACATCATTCCCAACGGGAAATCTGAAGACAGCATTAAGGGCCTTCAGCCTGAAGA
TGTTAATCTGTTGGTAACCTGCAGAGAGGAGGGCAACCCAGTGGCATCCTTAGGTCAITTCAGAGGCAGAAACACAGTTGACGC
TTACTCCAATGGGCAAGCAGCTCTCATGGGTGAGAAAGAGAGAAAGAAACGGATGATGGAGGTGCGTACTGGAAGTTTCAT
25 AGACTGGTTTGTGGCTTTAAAAGTAAGAGCCTCAGCAAGAGGAGTCTCAGAGACCTGATGGAAGAGGAGGCTGTTGTTTACA
GATGCTAGAAGAGACTCGGCAAGTTAAAGTAATAATAATTTGGACTTTTGTCTGTGTGTTCACTTGGAAATTTTCATGTTTGT
TTATTTCTCCTTATGAACCTAAGGATATGGTGAGACCTAACCTAAGACAATACTGACTGGTCTTTGGGAAAAAAGTTATGTA
ACTGTGCATCTCTCAGGCATTGTTTACGCTGTAGGTTTTAGCCAAATTTTACTTAGCAGAAAAATCATCTAATTTACAAGACTTTA
TTTCCAGAGATGGATTAAAGTAAATCTTCAACTTAAGTGAAGCCAAACCTAACAGACTGAATGTGCAAAATGTGGTTTTAA
30 TTTTGCATACCAAAGTAAGAAGAGACCAATTATTTCTCAGAGCACTTAGAGCAGAAATATATGTTAAGTTACCATGAATTAAG
TATACTGTCTGCACTGCCAAGTCTTGGCAGACCTTACCTGAAGTAGAAGATTGCTCATTTCTAAAGTTTTTTTCTGTCTCT
GTAATCCCTCCTACCATTAAAGAAAACTTATTTCTTAGACATTGTACAATCAGTTATGTACTGAAAATCGAATGTGCTGTGTG
ATACTTGTTCAGGACAAGTTTCATTTGCCAGGTTTCATTTGTAGCATGAGCCTACGGATTCTGATTTCCCAAGAAAGAAATGT
TTTCTGTAGGTATTTTGTACCACAGTATATGGAATGTTAGGAAAAAATTTGTTCCAGTTCCCTTTTTTTTTTTCTTTCTAC
35 TTTCAAGTTAAGTGAACCATACTGAAATGACCAACAAGTCTGCCTGTAAAGTTACATGTATGATTGTGTGTTAAATGATTA
TGGGGGAGAAAAAGTAAGTAAATGTGCTGATGATCCCATATTTATTGATCATATTAAGGTTGTTTATATAGTTTGGAAATGAC
CAGCCCCCTAAGCAGTGTTTGATTAACTTATGCTAATCAGATGATTACTCATATATTCTGCTAATTTCTAGCTTTATCTTGT
TATTTGGAAAAATATTAGCCAAATGCCCTTCTAGGTGGATCCAGTTGGAAGATATGTCAGAAACCTGAAGAAAAATGACGC
TGCCTTTGTGTGCTGGATTGCTCTACTTGATTAGATCATGATATATCAAGTTGAATTTTAGAGGGAAAAATTAATTTCTGATA
40 TCTTATTCATCCTTGATAAGTTTTTCCCTGATTTTTTTTTCTCCAAAGACTTTCCATCTGGACACAGCCTCCACATTTTTGT
TGATGTACTTAGAGCATAAGGATGTTTCAGTGGCAATTCGGCCGTGCGTAACAGAAAACTCAGTGCATACCTTTGCTGTGTT
AGGTTGTCAATATAGTCTTTCTGTAGGATGGATAGCATGTTTGGAGGTGCCAAACAAGAACTTTTGGGGTTAGTAGTGTGTCT
TGTGGAGGGTATTACAGGACTGTGTAATTATAGGACTCTAACCTGACATGGCTTGGCACCACCTTGCAGCTAGTGGGTACAGGG
TACAAAAGATGTTAGAGAAAAGCTCTACAGATTACGTACTTCTGTGCTTCTGATGCTCAACACTGTCTTTTGCTCTCATGAA
45 AGATGAAGGAAGCAAAATATGTAATGTAATTTTACCTTTGACCTTCTTAATCTCTGTATCTTTTAGATTGATGATTTTACTAGGC
TTGTATTTAGGGAATTTACTTTTCAATAATTTTGTAGATTTTGAATCAAACTCAGTCTTTTAAATTTTTTGTAGTCTATA
AACTAGTTTCATTATGATGGACTTGATTAGTCCAAAGTTAATTTTAGAAATTTGTCAGGTAGCATAGTCTTCCCATGATCAGG
AGGCTTTCTGAAGGACTGAGTCTGTAAATGAAAAAATAATTTATGTAATGAAATGATGATTTTCTGAAGAGCTTAGAGTGCCTC
GTAGAATTTTCTCAATTTTATTTCTGAGGTTTATAATTTGGGGGCCAAATAGATAGAGCTCATCATTTTCTTGTTTGGAAGTT
50 GAGGCTGCGACATGTCCAAGGTTATGAAGTCTCTTTTGGGAAGAACAGAAACAGGCTCCAAATCTGGACTCATGGTTTGTTC
AGATGTGTCTGGACAAATGGTTGTCAATGTTTGTCTGTTTTTCAAGGAAGTGTCTTCTTTGGGACAACTTTTGGTGT
TTGGGAAAGTAATAAGATCTTGGATTTTTCAAAATTAACATTAAGTTGTAAGAACTAAATTTTCTTTGAACCACTACTGTGT

AATTCAC TGATAATTGACATATTGGCTGGGCAGCCTATCTCTTCCATATCCAGCGTAAATGAATAGGAGGTGCTGTGATTTTT
TTTTTCTCCCTTTATTTAACATTGAGTCTAGTAGTTTGGAGAATTAGGGTCCCTCTACCTTCTTTCTGCTCTTGTCTTAGTAA
GATACATAAGGTACATCATCTTGTGCTGTGTATATAGCAGTAGGTCAAGTTTAGAGTACTAAAGTCTGTAAATAAGGAATG
ACTATTAGCATATTCATTAGAATTGTTTATTCTTGCCAGTATAAACATCATTTTATTTAGACTAAAGTCCCTGAGCCTGTCTT
5 TCTTATTGCTTTCCAGTAAATAGATAATGTGCTCGAGTAAGTTTGTGAATTGCTGATTGCAACTTAATTCAGGACCAGTCTTC
AATCTATATTTTCATTAGAATGATTGTCCCTGGAATGATCATACATGGACTGTCTTAAGCTAGCAAAATGTTTACTACTTTACACT
GACTAAATGGGTCCATAATGATGACATTGGTCTTTAGACATTAACATGTGTATATTTTTATATTAGCTCAAGCTAAGGTTTACAGA
ATTGAAGCTTGATATTGACTAGAAATAGCTAAAGTCAAAATGAGGTGAGGACACTGGTCTTGGGAAGGTAGAGAAAAATAAATGT
CTTACCAGGTGTTAATGGTATCCCCAGTTCTTAGACTTTTGTCTTCTCAGGCAATTTTCATCTCAAGATCTGATGAGAAGGGCA
10 TATTACATTGGTATGCAGGATGATTATTGCATATTTTGTGGGACCTCTAATTTCCCTGGTCACTTTTCAGAATATTCTGTTCTG
CCACCCCGAGAGAGTAAACACTTGAGCCGATTCTTCTTCCCGAGCTATTCTTTCTGGGGTAATTATGCTTTGTCTTTAGAT
TAGAGAAGCATCAAGCAATAGCAATGGTGTCTGTGCTTTCGGCCTAAATTCATAGATCTCATCTCCTAGGGCTTCTTTTCAC
TTGGCTCAAAGGATCCATTGTATTTTGGCACAAGAGCCTGGCCAGGGTCATGTAGCCATAGCTCTTAGGGATGATACCTCAAG
AAATTAGCTGGGACCCATCACTCTGTGAACTTCACATTTTAAAGAACTGAGTTGAGGGGTTGTTATGCATCTCTGTAACCTTGA
15 GGCTAAGCAAGGGTTAACTCTTGTGAGAGCCAAATAGAGTGTGTCTGTATTTCGAGTCCATGGCTCATTTTCTTTATAGTAGGC
ATATGGATCTTCCCTCTGACTTTGAATATCATTTTGGTGTGGCTGTGGGTATTTCATTCTTTACCACCAAATAAAGCGGCT
TATTAGCTACTCAGTTACTTGTACTCAAAGGTTAGGCTTCCCTGTTCTGCTTGGCAGTGTTAAAGCTTACAGGGTTAACTT
ATGATGATTCTCTGGCTCATTTTCATCAGAGGCATGATGACTGGAAGGGATCACATGGGTCGTTGGTGGTGACACCTCACTG
TTTCTTAGGTTTGGATAGAGAGATGTATACAAGACCTTTCTGTGTTAAATTCAGTACTACAGAGACTTGGCAGGACAAAATTTT
20 CTAAGAAATCAGAAAAATGATTAAAGTGAGATAAGTACCTGGGTGACACAGATATTAGCCCGTTGGTAAAGACAACAAATATTA
GCTTAAATCTGCATATGTAGAATCATTTTCATTAGATTAGAGCTTGAAGCACCTTGGCTCTCAGCTACTTTAACTCCTCCC
CATATAAATCAGGGCACCAATAAATAAGTTTCAGCTTTTAAACCTTGGTTGTATGTTAAGCATTATAAAGTACGAAGTTGT
ACCACAGTAGAGATAATTTAGTAGAAAAATGCTTTGAGGCTTCAGTATTGTGAAGATTTTGCATTAGCCAGATGCTAGGTTGTT
GAAGGCATTTTCAGTGTGATAATAGCCTGAGCAGACTTCTTTACAAATGGGATCTGTTTCTATATGTGTATATGCCACTTACC
25 ATTCAGAGAGACTGGTCTTTCTCTTTGTCTCCCTTCACATTGCTGTGTGCTAGTCTTACACCTAGTCTTTTTCAGCACTTAGCAAA
TCAAATTTGGATTTTTTGGTCAGCTTAGTTTCACTTTAAGGCATATTGGCATGGTGTGTGAAAGTGATGTTTGGCCCCAGTATTG
AGGACTTTTAGATCCAAATAATGACTCATTAATATAATATGTTTTAAGTATACTGAATTTCTGTTAGCTTAAATGTTAATC
CTCAGGAATGATTTCTTCACACTTTGTGTTGGCTAATAATAAAGCACTGTTTTATCTTCAAACTCTTTTTCAAAAATTAGG
GAGAGAGCAGTAGTGATCATTTATGTGAGCCCCTTTGAATGATGGTGTGAGAGTGCAGAGAAACAATGGAGTTTGTATGCCAA
30 AAAGGTTTTTTTGCAGTAAAGTAAAAATTTGGAATTAGTTGGCATATAAAGGAACCTTTTGTACTGGAACGTATGAGGCTGG
ATTGTGAAAAGGTAATCTTTCCATTGCTAGACTTGGTTAACTTAGGGCTGCAAACTCTTTTCTTCTATCAAGGTCACTTAATAT
GGAATGTTTTTGTGACAGCTGTCTTTGTTGGAATACTTTTAGCTGTTTACGCTACTTTGACTCCTAGGAGAGAATTTAGTTAAGGT
TCAAAGTAATTAACCTGGCTTTGCCAGTGGTGAGTCCACACCAATTTATTCACTTAGTAGTCATATAAATGTTTTTATTTTCTCTC
TTCAATGCTGAGAATAAGGCTTTAAATTAAGTATGATTCACCTTTAAAGGAATGTTGTGAGAATTGATGTAATTTCTGTTTCTGTTT
35 CCATCTAACTTCTTTATAAAGAGGGATAGTTTTTTTGTGTTTGGGGTAAGCACCTAATTTATCCAGTAACCAACAACCTTA
ACCATTTGGCATATATAGTCTTTCACTCAGAAATAAACAACCTGTTTGGTATATCTGTATCATTTGCTAATCTTGTGCACTTTA
CTTTTGGGCGTAGCCATACATAGTCTGAGGCTATTGACTTAAACCAATAACTGTACTTTATGTAATGACTCTTAAATTTGGTT
ACCTGGGTTTCACAGCTTGCTTGAAGAGAAAGGATGCTAGAATAAAGTAAGCAGCTGAAGAGCGAGCAATCAAGACAAAACACA
GTGGTCTCAGATTTTTCTGTAGTGTGGGAACAGTGGTTTTGCTCTATACCACTGAAAAGCACTATAACATAATTGTTGTCCATGA
40 TACTGAAGCTTTTCCCTCACTTCTAGGTTGTTTACATTAGAGCTCTATCAATAAGAGGAATACATATTACAGTGAATTCGAC
AACCACACAAGTTGGCAGTAGGTATCCCCAACCTAATTTATCTTGGTAAATTCACCTGTTTCCTAGTGTCTGGATAAAAGA
GTGTTTACTTTTTATTGCTCTTAGACAGAGTAGTCTAGATAAGTTTCAATTTATCAACATAGCCTAGACTTCTGTAAGTGGA
TGTTCAATTAGTAACATCATCTTTTGTGTTATAATTGGAACAGAAACGAGGCTTATTGCTATTGCAGAAATCCCAACTGGCA
AAGGCCAGTATATATGGTATTCCATAATAAACCAGCTTTTGAATTTATGTGTTGGATTAGTGCTTCTGTTTACCAGTATT
45 GACTCTGCTAGTTTGCACCTTTCCGTTCTTAAACAGAAAAATTTGTATTTGTTATTCCTCTTAAATTTTGTGCTAAGTGAAGG
AAGTAAAAAAGAAAAAATGATGATTTACATTGACATACTTTATGTGCAAGCTTTATTTAGGTTTCAAGTGAAGGAGGAGT
CTGTATTTGTGTTGATAGCTTAAATGTTGTTCTTTTATATCCATTAAAACTTAAAGTTACTTATGTTCTGTGATCTTAATTTT
GTTGTGTTTCCATTGTAGGTTGATAGGTATATCGAGAACAGGTACGTGACAACAGTTTATTTCCATGATAGAAAGCTAAAGTC
CATAGAAGCACAATAATCGTGTTCACACATTTAGTGTACCAACACATAGAAAGCACAAGACTAATAGTATTCTCTGTATCCACACA
50 AGTGCCAGTCATAAAGGCCACCAGGTATTGCTCAGAGTTGCTATGAGCACTACAGTATTGATAAGCCCAAGCAATGCCGTA
TCTAAACTGGTCTTAATGGTAAGGGACCCAAAGGAATAATCTCAATAAGTTTGTACCACATTGATGGAGGGAGAGAATAACAAAT
GTCAAGAATGCCAAATTTATATTGGGGTTACTAGCTAAATGGGGTTTGGAGGCTTTTACTGCAACTTGAAGTGGAGAAA

TAGGGACAGATGCTAGGTTTTTGGTGGGTGGAACAGGTGACATATTTCTGTTTAAAGCTGTAGTGTGATTGGGGTTTTTGTGTA
AAAAATCTTAAATCTTTTAGGAAATATTACCTCTTAACAGTGCCCCCAACATGCAGAAAGTCATACTTTAACAGGGCAAAT
ACTACTTGCTTTGATTTTTTTTGTGTACGTTTGTATGTGAGAGATGAAGTTACCTTTATTTTTTCTATACTTGACTGTGCT
TCATTTTAATAAAGGATAATTTGATCTGA

5 309

MSLQRLQHSSNGNLADFCAGPAYSSYSTLTGSLTMNDNRRIQMLADTVATLPRGRKQLALTRSSSLSDFSWSQRKLVTVEKQD
NETFGFEIQSYRPNQONACSSSEMFTLICKIQEDSPAHCAGLQAGDVLANINGVSTEGFTYKQVVDLIRSSGNLLTIETLNGTMI
LKRTLEAKLQVLKQTLKQKWEYRSLQLEHRLHLHGDAANCPSENMDLDELSLFGPLPGPGPALVDRNRLSSESSCKSWLSS
MTMDSSEDGYQTCVSEDSSRGAFSRQSTSTDDCEFIPEKGGDFLRRSSSRNRISINTSSGSMSPLWEGNLSSMFGTLPRKSRKGS

10 VRKQLLKFIPLHRAVEEESRF

310

GCACGAGGCGAGCTGGAGTGATCTTTATTCACAATGTCTTTACAAAGGCTCCTGCAACACAGCAGCAATGGCAATTTGGCGGAC
TTCTGCGCTGGGCCAGCGTATAGCTCTTACTCCACACTACCGGCAGCCTTACGATGAACGATAATAGAAGGATTCAAATGCTA
GCAGACACGGTGGCTACTCTGCCCGGGACGAAAGCAGCTTGCTTTGACCAGATCAAGTTCTTTAAGTGAATTTCTCGGTCT
15 CAAAGAAAGCTTGTTACTGTGGAGAAGCAGGATAATGAAACATTTGGATTTGAAATTCAGTCTTACAGGCCCCAGAATCAGAAT
GCCTGCTCCTCGGAAATGTTTACCTTTGATATGCAAAATACAGGAGGACAGCCAGCTCACTGTGCTGGCTGCAAGCTGGTGAT
GTCTTTGCAAAATATCAATGGTGTGAGCACAGAAGTTTTACCTACAAACAAGTCGTTGACCTGATCAGATCGTCCGAAACCTG
CTAACGATAGAGACTCTTAATGGAACAATGATTCGAAAAGAACGGAGCTTGAAGCAAAGCTGCAGGTTTTAAAGCAAACCTTTG
AAACAAAAATGGGTGGAGTACAGATCTCTGCAGTTACAGGAACATCGTCTGCTTCATGGTGATGCAGCTAATGGCCCCAGTTTG
20 GAAACATGGACTTGGATGAATTGTCTTTGTTTGGACCCCTGCCCTGGGCCAGGCCAGCCCTTGTGGACCGGAATCGATTATCC
AGTGAGAGCAGCTGTAAGAGCTGGCTGAGCTCCATGACGATGGACAGTGAAGATGGCTACCAGACGTGTGTCTGAGGACTCC
AGCAGGGGTGCCCTTCAGTCGGCAGACGAGTACAGATGATGAGTGCTTTATCCCCAAGGAGGGGGATGATTTCTGAGGAGGTCA
TCTTCAAGGAGGAACCGGAGCATCAGTAACACCAGCAGCGGATCCATGTCTCCCTTGTGGGAGGGCAACTTATCAAGCATGTTT
GGGACCTTGCCCGGAAGAGCAGAAAGGGAAGTGCCGAAAGCAACTCTTGAAATTTATCCCTGGCCTTCATCGTGTGTGGAA
25 GAGGAAGAAAGTCGCTTTTGACGGATTGTGGTGTCTTTCAAATTAGCTTATTTACAAATATCTCTAGACTACCCAGATCCC
AGCTTGGTGGGAAAGTGCAGAAGAAATGCAAAACTGACATCCCATTTTACAGCAATAGTGACCTTTATTTAAATTTGTTGTGTTA
TAGTTTATGCTTCTTAAATCATTTTTCAACCTAAACAGCCAATTTCTAAGCAGACAGGAAAATAAATAAGTTAATTAATA
TAACAAAGATGCAGGTTCTTGCTCATTCAGTAATGTCTTTGAAAGCAAACTAATATTTATTTCTAGATTATCCCTGTGAAT
AATTGAGAATTTTGGAGTCAAGTATGAATAAAGGTGTGGCAGAAATAAATACTGGACTATTTCTATAGGATAATTGCTG
30 GGTATAAAATCTTAGGTTTGCTTATGCCAGTAGCTCCTGCGGAGGCTTAATAATAGGCAATTTTGAATTTGTCAAACCTGT
AATGGCTTGTAACAAAGATGACCATCAGCTGTTTCTCACATCTATAGTGACAATAAAGCGGGAAGTATAAGATTTAATAGGAG
GGGTTAAGGTTTATGAGAACCATGGAAGATGTGGTCTGAGATGGGTGCTGCAAGATCATAATAAAGTCATTTTTATAGACAG
TCTAAACAAAAA

311

35 MEQQDQSMKEGRLTLVLALATLIAAFGSSFOYGYNVAAVNSPALLMQQFYNETYYGRTGEFMEFPLTLLWSVTVSMFPFGGFI
GSLLVGPLVNKFRGKALLFNIFSIVPAILMGCSRVSPELIIISRLLVGICAGVSSNVVPMYLGELAPKNLRGALGVVPQL
FITVGILVAQIFGLRNLNANVDGWPIILLGLTGVPAALQLLLLPFPESPRLYLLIQKKDEAAAKKALQTLRGWDSVDREVAEIRQ
EDEAEKAAGFISVLKLFMRSLRWQLLSIIVLMGGQQLSGVNAIYYADQIYLSAGVPEEHVQYVYTAGTGAVNVMTFCVAVFV
ELLGRLLLLLGFISICLIACCVLTAALALQDTVSWMPYISIVCVISYVIGHALGPSPIPALITEIFLQSSRPSAFMVGGSVHW
40 LSNFTVGLIPFPFIEGLGPYSFIVFAVICLLTTIYIPLIVPETKARTFIEINQIFTKMNKVSEVYPEKEELKELPPVTSEQ

312

CTTCTCTCTCCATTCACTGTCACGCGTTACTTTGGCTAAAGGAGGTGAGCGGCCTCTGCCCTTCCAGAGCAAGCATGGAGCAA
CAGGATCAGAGCATGAAGGAAGGAGGCTGACGCTTGTGCTTGCCCTGGCAACCTGATAGCTGCCCTTTGGGTATCCTTCCAG
TATGGGTACAACGTGGCTGCTGTCAACTCCCAGCAGTGTCTATGCAACAATTTTACAATGAGACTTACTATGGTAGGACCGGT
45 GAATTCATGGAAGACTTCCCCTTGACGTGTGCTGTGCTGTAAACCGTGTCCATGTTTCCATTTGGAGGGTTTATCGGATCCCTC
CTGGTCCGCCCCCTTGGTGAATAAATTTGGCAGAAAAGGGGCTTGTGTTCAACAACATATTTCTATCGTGCCTGCGATCTTA
ATGGGATGCAGCAGATCGCCACATCATTTGAGCTTATCATTTATTTCCAGACTTTTGGTGGGAATATGTGCAGGTGTATCTTCC
AACGTGGTCCCATGTACTTAGGGAGCTGGCCCCTAAAAACCTGCGGGGGCTCTCGGGGTGGTGGCCAGCTCTTCATCACT

5 GTTGGCATCCTTGTGGCCAGATCTTTGGTCTTCGGAATCTCCTTGCAAACGTAGATGGCTGGCCGATCCTGCTGGGGCTGACC
GGGGTCCCCGCGCGCTGCAGCTCCTTCTGCTGCCCTTCTTCCCCGAGAGCCCCAGGTACCTGCTGATTGAGAAGAAAGACGAA
GCGGCCGCCAAGAAAGCCCTACAGACGCTGCGCGGCTGGGACTCTGTGGACAGGGAGGTGGCCGAGATCCGGCAGGAGGATGAG
GCAGAGAAGGCCGCGGGCTTCATCTCCGTGCTGAAGCTGTTCCGGATGCGCTCGCTGCGCTGGCAGCTGCTGTCCATCATCGTC
10 CTCATGGGCGGCCAGAGCTGTCGGGCGTCAACGCTATCTACTACTACGCGGACCAGATCTACCTGAGCGCCGCGCTGCCGGAG
GAGCAGCTGCAGTACGTGACGGCCGGCACCGGGCCGTGAACGTGGTCATGACCTTCTGCGCCGTGTTCTGCTGGTGGAGCTCCTG
GGTCGGAGGCTGCTGCTGCTGGGCTTCTCCATCTGCCTCATAGCCTGCTGCGTGTCTACTGCAGCTCTGGCACTGCAGGAC
ACAGTGTCTGGATGCCATACATCAGCATCGTCTGTGTCATCTCTACGTCATAGGACATGCCCTCGGGCCAGTCCCATAACCC
GCGTGTCTCATCTAGATCTTCTTGCAGTCTCTCGGCCATCTGCCCTTCATGGTGGGGGGCAGTGTGCACTGGCTCTCCAAC
15 TTCAACCGTGGGCTTGATCTTCCCCTTCATCCAGGAGGGCTCGGGCCGTACAGCTTCATTGTCTTCGCCGTGATCTGCCCTCCTC
ACCACCATCTACATCTTCTTGATTTGTCCCGGAGACCAAGGCCAAGAGCTTCATAGAGATCAACCAGATTTTACCAAGATGAAT
AAGTGTCTGAAGTGTACCCGGAAGAGGAGAACTGAAAGAGCTTCCACCTGTCACTTCGGAACAGTGACTCTGGAGAGGAAGC
CAGTGGAGCTGGTCTGCCAGGGGCTTCCCACTTTGGCTTATTTTCTGACTTCTAGCTGTCTGTGAATATCCAGAAATAAACA
ACTCTGATGTGAATGCAGTCTCATCTCCAGCCTCCCCACCCAGTGGGAAGTGTGCAAAGGGCTGCCCTGCTGTTCTTGAAG
20 CTGGGCTGTCTCTCTCCATGTTGGCCTGTCAACAGACCCGAGTCAATTAACAGCTGGTCTCCACTTTGCTGGTTCAGCCTTC
GTGTGGCTCCTGGTAACGTGGCTCCACCTTGATGGGTCAACCTTTGTGTGGCTCCTGGTAACATAACAACAACAGTTACTATAG
TGGTGAGATGGAAGGAATCAAATTTTGCCAGAGAACTAACTCGGTGGCCCCAACAGGTCTTCCGGGGCCATGGGCATTTGTTT
AGAGCCAAATTCATCTCTTACCAGATCTTTTCCAGAAATACCTGTCTAGGAAGGTGTGATGTCAGAAACAATGACATCCAGA
AAGCTGAGGAACAGGTTCCTGTGGAGACACTGAGTCAGAATTCCTCATCCAATATTTTGTAGTGGAAATGGAATTGCTTC
25 TGTGTAGTCAATAAAATGAACCTGATCACTTTTC

313

25 MASVAVDPPQPSVTVRVNPLVSSYDLMSSAYLSTKDQYPYLKSVCEMAENGVKTTITSVAMTSALPIIQKLEPQIAVANTYAC
KGLDRIEERLPILNQPSQIVANAKGAVTGAKDAVTTTGTGAKDSVASTITGVMDKTKGAVTGSVEKTKSVVSGSINTVLGSRM
MQLVSSGVENALTKSELLVEQYLPTEEELEKEAKKVEGFDLVQKPSYYVRLGSLSTKLHSTRAYQALSRVKEAKQKSQQTISQ
LHSTVHLIEFARKNVYSANQKIQAQDKLYLSWVEWKRSIGYDDTDESHCAEQFESRTLAIRNLQQLQTTCHTLLSNIQGV
QNIQDQAKHMGVMAGDIYSVFRNAASFKEVSDSLTSSKQQLQMKESLDDVMDYLVNNTPLNLVGPFPQLTESQNAQDQGA
EMDKSSQETQRSEHKTH

313

30 MASVAVDPPQPSVTVRVNPLVSSYDLMSSAYLSTKDQYPYLKSVCEMAENGVKTTITSVAMTSALPIIQKLEPQIAVANTYAC
KGLDRIEERLPILNQPSQIVANAKGAVTGAKDAVTTTGTGAKDSVASTITGVMDKTKGAVTGSVEKTKSVVSGSINTVLGSRM
MQLVSSGVENALTKSELLVEQYLPTEEELEKEAKKVEGFDLVQKPSYYVRLGSLSTKLHSTRAYQALSRVKEAKQKSQQTISQ
LHSTVHLIEFARKNVYSANQKIQAQDKLYLSWVEWKRSIGYDDTDESHCAEQFESRTLAIRNLQQLQTTCHTLLSNIQGV
QNIQDQAKHMGVMAGDIYSVFRNAASFKEVSDSLTSSKQQLQMKESLDDVMDYLVNNTPLNLVGPFPQLTESQNAQDQGA
EMDKSSQETQRSEHKTH

35 314

40 ATGGCATCCGTTGCAGTTGATCCACAACCGAGTGTGGTGACTCGGGTGGTCAACCTGCCCTTGGTGAGCTCCACGTATGACCTC
ATGTCCTCAGCCTATCTCAGTACAAAGGACCAGTATCCCTACCTGAAGTCTGTGTGTGAGATGGCAGAGAACGGTGTGAAGACC
ATCACCTCCGTGGCCATGACCAGTGTCTGCCCCATCATCCAGAAGCTAGAGCCGCAAATGTCAGTTGCCAATACCTATGCCTGT
AAGGGGCTAGACAGGATTGAGGAGAGACTGCCTATTCTGAATCAGCCATCAACTCAGATTGTTGCCAATGCCAAAGCGCTGTG
45 ACTGGGGCAAAAGATGCTGTGACGACTACTGTGACTGGGGCCAAGGATTCTGTGGCCAGCAGATCACAGGGGTGATGGACAAG
ACCAAAGGGGCACTGACTGGCAGTGTGGAGAAGCAAGTCTGTGGTCACTGGCAGCATTAACACAGTCTTGGGGAGTCCGATG
ATGCAGCTCGTGAGCAGTGGCGTAGAAAATGCACTACCAAAATCAGAGCTGTTGGTAGAACAGTACCTCCCTCTCACTAGGAA
GAACTAGAAAAGAAGCAAAAAAGTTGAAGGATTTGATCTGGTTTCAAGCAAGTTATATGTTAGACTGGGATCCCTGTCT
ACCAAGCTTCACTCCCGTGCCTACCAGCAGGCTCTCAGCAGGGTTAAAGAAGCTAAGCAAAAAAGCCAACAGACCAATTTCTCAG
45 CTCCATTCTACTGTTCACTGATTGAATTTGCCAGGAAGAATGTGTATAGTGCCAATCAGAAAATTCAGGATGCTCAGGATAAG
CTCTACCTCTCATGGGTAGAGTGGAAAAGGAGCATTGGATATGATGATACTGATGAGTCCCACTGTGCTGAGCAATTTGAGTCA
CGTACTCTTGCAATTGCCCGCAACCTGACTCAGCAGCTCCAGACCAGTGCACACCCCTCCTGTCCAACATCAAGGTGTACCA
CAGAACATCCAAGATCAAGCCAAGCACATGGGGGTGATGGCAGGCGACATCTACTCAGTGTTCGCAATGCTGCCTCTTTAAA
GAAGTGTCTGACAGCCTCCTCACTTCTAGCAAGGGGAGCTGCAGAAAATGAAGGAATCTTTAGATGACGTGATGGATTATCTT

372

GTTAACAACACGCCCTCAACTGGCTGGTAGGTCCCTTTATCCTCAGCTGACTGAGTCTCAGAATGCTCAGGACCAAGGTGCA
GAGATGGACAAGAGCAGCCAGGAGACCCAGCGATCTGAGCATAAAACTCATTAA

315

5 MCLRGGCSPRAAAPQPRPPALPPRPRAPVPASRPGRPLLPARPCGRMRRGSPGPRLGGSRGERRRPAGRDPARVGFPGQGL
RRPARPGPAAWTETGQGIHALTDLSIPGMTSGNGNSASSIAGTAPQNGENKPPQAIIVKQILTHVIEGFVIEGADVSRWDAR
LLVGNLKKKYAQGFLEKLPQDDHTTTDSEMEEPYLOESKEEGAPLKLKCELCGRVDFAYKFKRSKRFCSMACAKRYNVGCTK
RVGLFHSRSLQKAGAATHNRRRPAKPVCHHLPRIPRSSQALCPFRLLLLCVTHSQEDSSRCSNDSYEEPLSPIASASSTS
AGDKASGTWSSPTCICGTWWAWDTTSCQVSHQVNVEDVVEFIRSLPGCQEIABEFRAQEIADGQALLLLKEDHLSVMNKLGA
LKIYARISMLKDS

10 316

GGCGCCGCATGTGTCTCCGCGGCGGCTGCAGCCCTCGAGCGCCCGCGCGCCCAACCCGCGCGCGCCCGCCCTCCCGC
CCCGGCTCGCGCCCCCGTCCCGGCTCGCGCCCGCGCCCTTTGTGACGCGCGCCAGGCCGTGCGGTGCGATGCGCCGCG
GCAGCCCCGGGCCCCGCTCGGAGGCTCCCGGGCGAGAGGAGCGCCCGCGCGCGGACCCGCGCGAGTCCGCCCCGGCC
AGGGGCTGCGTAGGCCCCCGGCCAGGCCAGCCGCTGGACAGAGACAGGGCAGGGCATTGTTCATGCACTGACCGACCTCA
15 GCATCCCCGGCATGACCTCAGGGAACGGAACCTGCGCTCCAGCATCGCCGGCACTGCCCCCAGAATGGTGAATAAACCAC
CACAGGCCATTGTGAAACCCAAATCCTGACGCATGTTATCGAAGGGTTGTGATCCAGGAGGGGCGGACGTTTCCCGGTGGG
ACGCTCGTCTGCTGCTGGGGAATCTCAAGAAGAAGTATGCACAGGGGTCTCTGCTGAGAACTCCACAGCAGGATCACACCA
CCACCACTGACTCGGAGATGGAGGAGCCCTATCTGCAAGAATCCAAAGAGGAGGGTGTCTCCCTCAAACCTCAAGTGTGAGCTCT
GTGGCCGGGTGGACTTTGCTTATAAGTTCAAGCGTTCCAAGCGCTTCTGTTCATGGCTTGTGCAAGAGGTACACGTTGGGAT
20 GCACCAAACGGGTGGGACTTTTCCACTCAGACCGGAGCAAGCTGCAGAAGGCAGGAGCTGCGACCCACAACCGCGCTCGGCCAG
CAAAGCCAGTCTGCCACCCTTACCAAGGATACCAAGAAGCAGCCAACAGGCACTGTGCCCCCTTTCGGTTACTGCTGCTTTGCG
TAACACACAGCCAGGAAGACTCCAGCCGTGCTCAGATAACTCAAGCTATGAGGAACCCCTGTACCCATCTCAGCCAGCTCAT
CTACTTCGCGCGCGACAAGGCCAGCGGGACCTGGAGCTCCCGACATGCATATGCGGGACCTGGTGGGCATGGGACACCACTT
CCTGCCAAGTGAGCCACCAAGTGAATGTAGAAGACGTCTACGAATTCATCCGCTCTCTGCCAGGCTGCCAGGAGATAGCAGAGG
25 AATTCGCTGCCAGGAAATCGACGGGCAAGCCCTGCTGCTGCTCAAGGAGGACCACCTGATGAGCGTTATGAACATCAAGCTGG
GGCCGCGCTGAAGATCTACGCCGCATCAGCATGCTCAAGGACTCCTAGGGCTGGTGGCACCAGGATTCTGGCCAGGGCGCC
TCCTCCCGACTGAGCAGAGCCAGACAGACATTCCTGAGGGGCCAGAAATGGCGCGCTTGGAGGGCAGGGGCTCTCCCTAGGGG
CATAGCTGGTGAGGAGGTCTGGGCACCTCTCCATGGCTCTCAGGGGCTTTCAATTCTGTGGGAGGGGAGAGAGGTAGGTGG
CACAGAAGATGGGGCTTTATGCTTGTAAATATTGATAGCACTGGCTTCCCTCCAAAGTCCCAATACTCTAGCCCCGCTCTCTTCC
30 CCTCTTTCTGTCCCCCATTTTCCAGGGGTATATGGTCAGGGCTCCCCAACCTGAGTTGGTTACTTCAAGGGCAGCCAGCAGGC
CTGGATGGAGGCTAGAAAGCCCTTGCTTCTTCTCCCACTTCTTCTCCAGGCTGGTTAACTCTTCCGTTGTCAGCTTCT
CCCCCTTCCAGCTGTTTCTGACAGAGCCAGGGTCTCCCCCTACACCTCTGCAGGTGGAGAGAGAGAAGCTGGGCCAGCCG
CGGTGCTGCTGGCAAGACGCTTAACGCTGTGTATGACTGTGTGACTGTGTGGGAGCCTGGACTGACAGATAGGCCAAGG
GCTACTCTCTGGCATCTCCAGGTGTTTGTAGCAAAACAGCCACTTAGTGCTTTGTCTGACTCCACTCAGCCTCAGGATGGGG
35 AATAGCCAAGAAATGGCAGCCTCAGCGCAGAGGCAAGGTGAGAAAGAGACGGCGCTTCAGAGTTTCTTTCCAGACACCCCTCCC
CGCACTGTGAAGTTCCCTTGACCGCCTCTGCTTCAAAAGAGCATTAAGAAAGCTGCGGTGGTCTGAGCAACATAGCCGAGA
CGTGGAGCCTCTGCGCTGCTGCCCCGCCACCTGGGAGTCCAGTGGTGAGGCTCAGAGAACTTCAAGGGGAAAGAACAGCT
GGAGTTCTGTGTGATGTGAAGAAGGCAGCTTTGGCCTCCCACTCCCACTTCTTTGCTTATAAATCTTCTAGCAGCAATTT
GAGCTACCTGAGGAGGAGGCAGGGCAGAAGGGCAAGGGCTGCCTCTGACCTGCGGTGTCTTTGAGGAAAGGAGGTAGGCACC
40 TTTCTGAGCTTATTCTATTCCCCACCCACACCCAGGCAGGGTTGGAATGAAGGACTTTTTTAACCTTTGTTTTGTTTTTA
AAAATAAATCTGTAAATCTGAAAAAAAAAAAAA

317

5 MHPFYTRAATMIGEIAAAVSFISKFLRTKGLTSEQLQTFQSLSQELLAEHYKHHWFPEKPKGSGYRCIRINHMDPLIGQAA
45 QRIGLSSQELFRLLPSELTLWVDPYEVSYRIGEDGSICVLYEASPAAGSTQNSTNVQMVDSSRISCKEELLGRTSPSKNYNMMT
VSG

318

CCTCTCGGAGCTGGAAATGCAGCTATTGAGATCTTCCAATGCTGCGGAGCTGGAGGCGGAGGCAGCTGGGGAGGTCCGAGCGAT
GTGACCAGGCCGCCATCGCTCGTCTCTTCTCTCTCTGCGCCCTCCTGTGTCGAAAATAACTTTTTTAGTCTAAAGAAAGAAA
GACAAAAGTAGTCGTCCGCCCTCACGCCCTCTTCTCTCTCAGCCTTCCGCCCGGTGAGGAAGCCCGGGGTGGCTGCTCCGCC
5 GTCGGGCCCGCCGCGAGCCCCAGCGCCCCGGGCCGCCCGCACGCCGCCCATGCATCCCTTCTACACCCGGGCCGCCA
CCATGATAGGCGAGATCGCCGCCCGGTGTCTTCTCTCCAAGTTTCTCCGCACCAAGGGGTGACGAGCGAGCGACAGCTGC
AGACCTTCAGCCAGAGCTGCAGGAGCTGCTGGCAGAACATTATAAACATCACTGGTTCAGAAAAGCCATGCAAGGGATCGG
GTTACCGTTGTATTCGCATCAACCATAAAATGGATCCTCTGATTGGACAGGCAGCAGCGGATTGGACTGAGCAGTCAGGAGC
TGTTACGGCTTCTCCCAAGTGAACCTCACACTCTGGGTGACCCCTATGAAGTGTCTACAGAATTGGAGAGGATGGCTCCATCT
GTGTGCTGTATGAAGCTCACCAGCAGGAGGTAGCACTCAAAAACAGCACCAACGTGCAATGGTAGACAGCCGAATCAGCTGTA
10 AGGAGGAACCTCTCTTGGGCAGAACGAGCCCTTCCAAAACTACAATATGATGACTGTATCAGGTTAAGATATAGTCTGTGGAT
GGATCATCTGATGATGATCCATAAATTTGATTTTTGCTTTGGGTGGGCTCCTCTTGGGGATGGATTATGGAATTTAAACCATGT
CACAGCTGTGAAGATCTGGCACAAGATAGAATGGTAAAAAATAAATTTAAGTGACAGTGCCATAGTTTGGACAGTACC
TTTCAATGATTAATTTAATAGCCTGTGAGTCCAAGTAAATGATCACTTTATTTGCTAGGGAGGGAAGTCTAGGGTGGTTTCA
GTTTCTCCAGACATACCTAAATTTTACATCAATCCTTTTAAAGAAATCTGTATTTCAAGAATCTTCTCTGCAGTAAATC
15 TCGCAGGGGAATTTGCACTATTACACTTGAAAGTTGTTATTTGTTAACCTTTTCGGCAGCTTTAATAGGAAAGTTAAACGTTTT
AAACATGGTAGTACTGGAAATTTTACAAGACTTTTACCTAGCACTTAAATATGTATAAATGTACATAAAGACAACTAGTAAGC
ATGACCTGGGGAAATGGTCAGACCTTGTATTGTGTTTTTGGCCTTGAAAGTAGCAAGTGACCAGAATCTGCCATGGCAACAGGC
TTTAAAAAGACCTTTAAAAAGACACTGTCTCAACTGTGGTGTAGCACCAGCCAGCTCTCTGTACATTTGTAGCTTTGTAGTT
TTCTAAGACTGAGTAAACTTCTTATTTTAGAAAAGTGGAGGTCTGGTTTGTAACTTTCTTGTACTTAATTGGGTAAAGTCTT
20 TTCCACAAACCACCATCTATTTTGTGAATTTGTTAGTCACTTTTATTTGGTAAATATGAAGTGGTGAATTTGTACAGTT
CATGTATATTGATTGTGGCAAAGTTGTACAGATTTCTATATTTTGGATGAGAAATTTTCTCTCTCTATAATAAATCGTTTCT
TATCTTGGCATTTTTAACC

319

MATKARVMYDFAEPGNNELTVNEGEIITITNPVGGWLEGRNIKGERGLVPTDYVEILPSDGKQDFSCGNSVADQAFDLSLS
25 ASTAQSSSSAASNHNQVSGNDPWSAWSASKSGNWESEGWGAQPEGAGAQRNTNTPNNWDTAFGHPQAYQGAPATGDDDDWDED
WDGPKSSSYFKDSESDAGGAQRNSRASSSMKIPLNKFPFGAKFQTEQYLLAKQLAKPKKEKIPPIVGDYGPWVYPTSTPDC
VVADPRKGSKMYGLKSYIEYQLTPTNTNRSVNHRYKHFDWLYERLLVKFSAIPIPSLPDKQVTRGFEEFIMRMRERLQAWMT
RMCRRHPVISESEVFQQLNFRDEKEWKTGKRKAERDELAVMIFSTMEPEAPDLDLVEIEQKCEAVGKFTKAMDGKVELLTVG
QEHWRCTGPLPKEYQKIGKALQSLATVFSSSGYQGETDLNDAITEAGKTYEEIASLVAEQPKDLHLMECNHEYKGLGCFP
30 DIIGTHKGAIEKVKESDKLVATSKITLQDKQNMVKRVSIMSYALQAE MNHFHSNRIYDYNVIRLYLEQQVQFYETIAEKLQAL
LSRFPVM

320

CGGGAGACGAGCCGGCCGTCCCGGGCCGGGGACCCGCCCATGGCCACCAAGGCTCGGGTTATGTATGATTTTGTCTGCTGA
ACCTGGAAATATGAACGACGGTTAATGAAGGAGAAATCATCACAATCACAATCCGGATGTAGGTGGAGGATGGCTGGAAGG
35 AAGAAACATCAAAGGAGAACGAGGGCTGGTTCACACAGACTACGTTGAAATTTTACCCAGTGATGGAAGATCAATTTCTTG
TGGAAATTCAGTGGCTGACCAAGCCTTCTCTGATTCTCTCTCAGCCAGCAGCTCAGGCCAGTTCTGTCGGCTGCCAGCAACAA
TCACCAGGTTGGCAGTGGCAATGACCCCTGGTCAGCTGGAGTGCTCCAAATCTGGGAACCTGGGAAAGCTCAGAAGGCTGGGG
GGCCAGCCAGAGGGGGCTGGAGCCCAAAGAAACACAAACACTCCCAACAACTGGGACACTGCCTTCGGCCACCCCAAGGCTTA
CCAAGGACCAGCAACTGGTGATGATGATGACTGGGATGAAGACTGGGATGGGCCCCAAATCCTCTTCTACTTTAAGGATTCAGA
40 GTCAGCTGATGCAGGCGGCTCAGCGAGGAAACAGTCGTGCTAGTTCTCTCATCCATGAAAAATTCCTTAAACAAATTTCTCTGG
ATTTGCGAAACCTGGCACGGAACAGTATTTGTTGCCCCAAACACTAGCAAAACCCAAAGAGAAAAATTTCCATCATTTGTGGAGA
TTATGGCCCAATGTGGGTTTATCCTACCTCTACTTTTGAAGTGTGTGGTAGCAGATCCCAGAAAAGGCTCCAAAATGTATGGTCT
AAAGAGCTACATCGAATATCAGCTAACACCTACTAACACTAATCGATCTGTAAACCACAGGTATAAGCACTTTGACTGGTTATA
TGAGCGTCTCTGGTTAAGTTTGGGTGAGCAATTCATCCTCTCTTCCAGACAAACAAGTCACAGGCCGCTTTGAAGAGGA
45 ATTTATCAAAATGCGCATGGAGAGACTTCAGGCTGGATGACCAGGATGTGTGCCATCCAGTAATCTCAGAAAGTGAAGTTTT
CCAGCAGTTCTTAAATTTCCGAGATGAGAAGGAATGGAAGTGGAAAGAGGAGAGAGATGAGCTGGCGGGAGTCAT
GATATTTTCCACCATGGAACAGAGGCACCTGACTTGGACTTAGTAGAAATAGAGCAGAAAGTGGAGGCTGTGGGGAAGTTTCA
CAAGGCCATGGATGACGGCGTGAAGGAGCTGCTGACGTTGGGCAGGAGCACTGGAAGCGCTGCACGGGCCATTACCCAAGGA
ATATCAGAAGATAGGAAAGGCTTGCAGAGTTTGGCCACAGTGTTCAGTTCCAGTGGCTATCAAGGTGAAACAGATCTCAATGA

TGCAATAACAGAAGCAGGAAAGACTTATGAAGAAATTGCCAGTCTCGTGGCAGAACAGCCAAAGAAAGATCTCCATTTCTCGAT
GGAATGTAATCACGAGTATAAAGGTTTTCTTGGCTGCTTCCCTGACATCATTGGCACTCACAAGGAGCAATAGAAAAAGTGAA
AGAAAGTGACAACTAGTTGCAACAAGTAAAATCACCCACAAAGACAAACAGAACATGGTGAAGAGAGTCAGCATCATGCTTTA
CGCGTTGCAAGCTGAGATGAATCACTTTACAGTAACCGGATCTATGATTACAACAGTGTATCCGCCTGTACCTGGAGCAGCA
5 AGTGCAATTTTACGAAACGATTGCAGAAAAGCTGAGGCAGGCCCTCAGCCGCTTTCCAGTGATGTAGGACAGAACGGGCCCTTGA
AGAGAATGCCGCGTGCTTTCTCTGACTTGGGGCAATGCAATTCAAAACCTTTTTTCCCTATTATTTCAGAAAAAAGGAAAC
AAAACCAAAAAGAAAGAGTTGCAAAAACGCAATTTATTTATTTAGCCACCCTAAATGCGTCAGTTATTTAGGGATGGTCTTTT
GTTCAATTTCCGCATCCATTTATTTAAACAGTGGAAATTTGCTCTATTTTGGAAAGTACTTAAAAGTTACCAGAATTTTCAATG
GAAAATGAGGGGTTTCTCCCACTGATATTTACATAGAGTCATAATTTATATGCTCTTATAAATTATAAGTCTTATATAATTTA
10 TAAGTCTCCCAATCTTCCAGTTCTTACCCAGTGTGAGATAATTAATTACTAATTACTTTCTTAAAAACATGAACATATGCCAG
AATAAAAAATATCTATGTTTGTATAAAAAATAAAAAA

321

MKDLGAEHLAGHEGVQLLGLLVYLEQEERFQPREKGLSLIEATPENDNTLCPGLRNAKVEDLRSLANFFGSCSTETFFVLAVNIL
DRFLALMKVKPKHLSCIGVCSFLLAARIVEEDCNIPSTHDVIRISQCKCTASDIKMEKIISEKLHYEATTAALNHLHYHTI
15 ILCHTSERKEILSLDKLEAQLKACNRLIFSKAKPSVLALCLLNEVELKSVELLEILLVKKHKSINDTEFFYWRELVSKCL
AEYSSPECKPDLKKLVIVSRRTAQNHLNSYYSVPELPTIPEGGCFDESESEDSCEDMSCGEESLSSSPSDQECTFFFNFKV
AQTLCFPS

322

CCGGTCGGATGCCGACCGGGGGCACCGCTGAGGCGGTGGGTCCCCGACCTGCGAGACAGGTTTGGAAAGCCCCGCTGCGCCCA
20 GTCCGTGCGGACCGCGAGGCGCGGGCGGGTGGAGGCGGTCTCCGGCACGATGAAGGATTTGGGGGCAGAGCACTTGGCAGGT
CATGAAGGGGTCCAACTTCTCGGGTGTGTAACGCTCTACCTGGAACAAGAAGAGAGATTCCAACCTCGAGAAAAAGGGCTGAGT
TTGATTGAGGCTACCCCGGAGAATGATAACACTTTGTGTCCAGGATTGAGAAATGCCAAAGTTGAAGATTTAAGGAGTTTAGCC
AACTTTTTTGGATCTTGCACTGAAACTTTTGTCTGGCTGTCAATATTTTGGACAGGTTCTTGGCTCTTATGAAGGTGAAACCT
AAACATTTGTCTTGCACTGGAGTCTGTCTTTTTTGTCTGGCTGTGATAGTGAAGAAGACTGCAATATTTCCATCCACTCAT
25 GATGTGATCCGATTAGTCAGTGTAAATGTACTGCTTCTGACATAAAACGGATGGAAAAATAATTCAGAAAAATGCACTAT
GAAATGGAAGCTACTACTGCCTTAAACTTTTTGCACTTATACCATACTATTATCTTTGTCTACTTCAGAAAGGAAAGAAATA
CTGAGCCTTGATAAACTAGAAGCTCAGCTGAAAGCTTGCAACTGCCGACTCATCTTTTCAAAGCAAACCATCTGTATTAGCC
TTGTGCCTTCTCAATTTGGAAGTGAAACTTTGAAATCTGTTGAATTAAGTGAATTTCTCTGTAGTTAAAAACATTTCCAAG
ATTAATGACACTGAGTTCTTCTACTGGAGAGAGTTGGTTTCTAAATGCCTAGCCGAGTATTCTTCTCTGAAATGTTGCAAAACCA
30 GATCTTAAGAAGTTGGTTTGGATCGTTTCAAGGCGCACAGCCAGAACCTCCACAACAGCTACTATAGTGTCTCTGAGCTGCCA
ACGATACCTGAGGGGGGTGTTTTGATGAAAGTGAAAGTGAGGACTCTTGTGAAGATATGAGTTGTGGAGAGGAGAGTCTCAGC
AGCTCTCTCCAGTGATCAAGAGTGCACCTTCTTTTTCAACTTCAAAGTGGCACAAACACTGTGCTTTCCATCTTAGAAATCT
GATGTCTCTGTCAGAATTTATATTTACAGGGTTTCAAAGCAATAAATGGGGAATAGGTAGTTTCTGGTTTAGCCCCCATCTA
GTCAGGAATTAATATACTGGAATACCTACCTTCTATTTGTTATTCAGATCAGATCTGGCCTATTTTCAATTTATCCTAAGCCA
35 TCAATGGGGTAGTGCTCTTAAACCATTAACAGTACTTTAGACATGGCACTTTATTTTCTCGTAGATCTTTAGCTACTTTG
GGGAGGAGGAAGGTGCTGATACCTTCAATTTGTTACTTTTCAAGATTTTAAAAATAACTAGTGTAGCTTATCTTAAACATTT
TATAAAACCTTCAGATGCTTTAAGCAGATTGGAAGTATGCAAGTCTTCTTAGCAGGGACAGTGGAATAATCCTTAATGGTTT
ATCATAGATTTACCCCTCCCCCTTCTCAGAAGAGTGAGTATGCTTAAATGTCAAACACATTTTGTGTTTGTGTTTTTAA
ATGATCAGTGTCTATTTGATGTGATGCAGATCTTATAAATTTGGGAATTATAATATTGACATTTCTGTGATTTTATATATGTA
40 ATGTCTTAATTGAGATTTGTGTTAAGGCAGAAATAATTAGGCTAGGGCTCTTAGTTTTCATTCCTATTGCCCAAGTATTGTCA
AACTATGGTATTATTTAATGTTACTTTAAAAATCCATAATCTGCTAGTTTTCATGTACTTATATGAAAACAGTGCAGTAAGT
TGAAAACCTCAGTATCTATGGAATTGATAAATGTTGATCTGGTGTAGTATATTTTATCCGCATTTTCTTATATTAATAAATGTTT
TGCATGATTACATTTTATTTGCCTTTGT

323

MPPHPLNKHHTHTHTHTHTHTHMLLGTAPSCSLRMAVQAORRAVGKTSWSTLSPHTSQALHCLAGILHGICHACSQGEKS
SCHSAKSKNPGKVYKMTLDPVKQVSPELAMQ

324

AGGCAGTGTGTGGCCAGGGAAAGCTGCGTTTCATGGAGGCTTCCAGGAGGCTGCCCTGTAAAGCCAGGCAGGTGGCTTCAAGT
GCCATTCTGCAAGTGCCCTGCTTCAGCTCTCAGGCCCATCCAATGCCCTCGAGCCACGGGGTGTCCCGGCAGGTCTGCGTC
ATCACGACCCAGTAACGCTGCTGGCTGTGCTCATGCCCTCTTCCTCTGGCTGGCCAGTCATCTTTGTCACTTCTACACCTTT
GAGTGTGTACAGGCTGAATGTTGAGTGTATGCCCCCACCCTCTCAATAAACACACACACACACACACACACACACACAC
5 AAAATGTTATTAGGCACAGCAGCTCCATCCTGCTCCCTGAGAATGGCAGTGCAGGCACAGAGAAGAGCAGTTGGGAAGCAGGC
AGCTGGTCTACCTGTCCCGGCACACTTCACAGGCACCTCAGTGTGGCAGGAATTCATGCATGGCATCTGCCACGCATGCTCA
CAGGGTGAAAAATCAAGCTGTCACTCTGCAAAATCTAAAAACCTGGAAAGGTCTATAAGATGACATTAGACCCAGTTAAGAAG
CAAGTGTACCTGAGTTGGCCATGCAATGAATGAATATTATAAGCAGTATTGGAGAACTAGGCGGTGCTTACTAAGAAGTGTCT
TTGGAACATTCGAGTTTATACAGTGCATGACAAAGTAATAGAAGGCTGCAAGCACTGAAAAATGCACCCGAATCTGGTACCTAG
10 TGAGTGTACATTACTGGTGATTGAAGTATTTTCAGGAGCAGATACATGTGGTGATGTTATATGCACTTAGTTTGAAGTATCCC
TCACCCAAAAAGTGCAGAACCTCTGAAAGGTGAGCCGTGGTGCAAGTGGCCCCCTCGCTGTTCCGGTCACTTCTGGCCATGCT
TGGCCTGTACACGCCGGGTCTCAACAACATCGTGCGAGGAGACAGTGAAGTGTTCAGCCCAAGAGTCAACTGTTGAGCATGA
CATCAGGAAGCCTGTACTGTCTGATCCACCAGGCTTGGGTAAATTTCTATTGCTTCAGAGAAAGAGAGGCGCTCTTGTGCT
15 TTTGAGAGAGCTCTCATGTTAAAAAGAGGTGTCTTTGCACTGCGCTTGGTATGATTTGTTTCTTATAAAACAGGAAGAGAAC
CTAGAATGTGCTTTTCGTGAAGGTCTGACAAATGGGGTCTCAGAGTCTTAGGACCCGCACCAAGGTCAGAGTCCACCTGTAGAG
GCTGAGCAITTTCTACATGCACCCAGGGTGTGGGCTGAGGCTAAGGGTGGACGCGTAGACCCATCACAGGCTTCTTCTCTGAG
CCCAAGTCCCCCATTTGGTCACTATCCCTTCTCAGAGCAAGTGGATTTCTTTTCTCTCACTTGAATATTCAATGTTGAAGTAG
TGGCCTGGAGTGGCCTTTCCCTCAGCTGCTGTCTGGGCTCCCTGGCTTCCCGATCTTCACTCTCATACGTGAGGCAATAGGAG
20 GAGTTTCTATTCAATTGAGACAAACAGTGAAGGGCACCACAGGCCAGCAGCTTGGCCCTTTTCACAGAGGTGGCCTTGGAGTC
TGGAGTCTGGGAGCTGTCTGCCTCTCTACGGCTGTGCTTCCCAGCGGCACACGGGTCCGGAGCTCAGAATGAAGTGGTGCA
GAGGCGCAGGGCACACCTAGGAGAAGCTCAATGACCACCAAGCCCTTGAGGAGAGACGGGGTTTGTGTGTTAGAGAAATGGC
AACCCCAAGGACCATGCACGGCTCTGGCAGTGCACAGTGGCTGATTTTACCAGTTCCTCCTACCCCTGCCTCTGTCGGAGACT
AGAAAAATTTCTACCCCTTCAGGTCTTCTTTTAAAAAGAAAGAGCTATGGCAAAGCGTGCCATACAAAGTTAAAAAAGGAAGA
25 GTGGCTTTGCGTTCTTGACCATGGCCACAATTTGGTCACTTCTAGCATCTGACCTTGAGAACGGTGTAGAGGCTCTGAGCCAC
CGCTGGCCTTTGAGAGAGTCTCACTCTCCAGCTAGCTCTTTGTGTAGTGTCCAGGGCCCGGCAGTCAGGATAGAGCGAGCAAA
AGAAGATGGACAGTGGACAGGGCGCAGTGGCTCAGCCCTGTAATCCCAGCGCTTTGGGAGAGCGAGGCGGGCAGATCACAAGGT
CAAGAGATCAAGACCATTCTGGCCAACATGGTGAAACCTGTTTCTACTAAAAATACAAAAATTAGCTGGGCATGGTGGTGCGT
GCCTGTAATCCAGCTACTCGGGAGGCTGAGGCAGGAGAAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGATCGT
30 GCCACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCGCTCTCAAAAAAGATGGCCATGCTGGCCAAGGCTGGGCTCATAGGAA
ATATCTGTAGTAGGATGAAAGGAATCTGGCTAGGGAGCCCCGACTCGGTTTACCTACCCCGTACTCATCCACTCATCTGTC
TACCTAGTGTCCATCATGCCCTGGATGAATCATAGGTGGGAGGGACAGAAGCCGAGAACACCAGGCACACGGGGTGGGGGAA
TGGGGGAGCCAGGTGATCCCTTTGGAAGTGGATTGATCTCTCCCTGTGGAATGACAGCTTAACCTGATTAACCCACTG
GGGCTTGCTATTTTATGAGCATTTTAAACAGAAAGGAAGATGATATCCAGTGCAGGTAGACTTTATTTTGAAGCATCGT
35 GGCCCTCGCCTCTCCCTCTCCCCAGAACCTTGGCCAGAATTGAGTCTTCGCGTTGGGGAGGCCACGCTGGAACCATTTGCTCT
AGGGCTGAAGGTGGAAGTAGGGCCACTGCCGAGCTTCAGCCAGAGCCCTCGCTGCTTCGAGAAATTAACAAGCACACGACTG
GGATTCCCTTTGGATATGAGAAAGGAAGGCAAGTGGGGCAAAGGGAAGCTCTCCGTTTTTCATAAACAGAAAGTAAGATGCTC
TCTTCCCAGGTTTGCAAAATCACTTCTGGCAAATTAATTGAAGGTCTTGAATTACAGCCAGTGCACGAAACCTTTGCATGTTCT
TTCTGAAACTGTACTTCTTCCCTATTTCTGTCTACCCACACTCTGCGAACTTTGCCTTTTCTTAAACACATTCCAGACCAACA
40 CTGTTCACTTGTGTTAAAAA

325

MGTSLSPNPDWPLNPLSIQQTLLLLLSVLATVHVGRLLRQRRRLRSAPPGPFAWPLIGNAAAVGQAAHLSFARLARRYGDV
FQIRLGSCPIVVLNGERAIHQALVQQGSFAFADRPAPAFSFRVVSGRSMFAGHYSEHWKVQRRRAHSMNRNFTTRQPRSRQVLEG
HVLSEARELVALLVGSADGAFLDPRPLTVVAVANVMSAVCFGCRYSHDDPEFRELLSHNEEFGRTVGAGSLVDVMPWLQYFPN
45 PVRTVREFEQLNRNFSNFILDKFLRHCESLRPGAAPRDMMDAFLLSAEKKAAGDSHGGGARLDLENVPATITDIFGASQDTLS
TALQWLLLLLTRYPDVQTRVQAEQDQVGRDRLPCMGDQPNLPYVLAFLYEAMRFSFVPVTFPHATTANTSVLGYHPIKDTVV
FVNQWSVNHDVPVKWPNPENFDPARFLDKDGLINKDLTSRVIMFSVGKRRICGEELSKMQLFLFISILAHQCDFRANPNPAKMN
PSYGLTIKPKSFKVNVTLRESMELLSAVQNLQAKETCQ

326

ACTCTGGAGTGGGAGTGGGAGCGAGCGCTTCTGCGACTCCAGTTGTGAGAGCCGCAAGGGCATGGGAATTGACGCCACTCACCG
ACCCCCAGTCTCAATCTCAACGCTGTGAGGAAACCTCGACTTTGCCAGGTCCCCAAGGGCAGCGGGGCTCGGCGAGCGAGGCAC
CCTTCTCCGTCCCCATCCCAATCCAAGCGCTCCTGGCACTGACGACGCCAAGAGACTCGAGTGGGAGTTAAAGCTTCCAGTGAG
GGCAGCAGGTGTCCAGGCCGGGCTTCCGGGTCTCTGTTGACGCTTTGCCCTAGGCCAAAGGTCCCAGTTCTTCTCGGAGCCGGC
5 TGTCCCGCGCCACTGGAAACCGCACCTCCCCGAGCATGGGCACCAGCCTCAGCCCCAAGACCCTTGGCCGCTAAACCCGCTG
TCCATCCAGCAGACCAACGCTCCTGCTACTCCTGTGCGGTGCTGGCCACTGTGCATGTGGGCCAGCGGCTGTGAGGCAACGGAGG
CGCAGCTCCGGTCCGCGCCCCCGGGCCGTTTGCCTGGCCACTGATCGGAAACGCGCGCGGTGGGCCAGCGGCTCACCTC
TCGTTTCGCTCGCCTGGCGCGGCGCTACGGCGACGTTTTCAGATCCGCTGGGCAGCTGCCCATAGTGGTGTGAATGGCGAG
CGCGCCATCCACCAGGCCCTGGTGCAGCAGGGCTCGGCCCTTCGCCGACCGGCCGGCCTTCGCCCTCCTTCGGTGTGGTTCGGC
10 GGCCGAGCATGGCTTTCGGCCACTACTCGGAGCACTGGAAGGTGCAGCGGCGCGCAGCCACAGCATGATGCGCAACTTCTTC
ACGCGCCAGCCGCGCAGCCGCCAAGTCTCGAGGGCCACGTGCTGAGCGAGGCGCGCAGCTGGTGGCGCTGCTGGTGCGGC
AGCGCGGACGGCGCCTTCTCGACCCGAGGCCGCTGACCGTCTGGCCGTGGCCAACTCATGAGTGGCGTGTGTTTCGGCTGC
CGCTACAGCCACGACGACCCCGAGTTCCGTGAGCTGCTCAGCCACAACGAAGAGTTCCGGCGCAGCGTGGGCGCGGGCAGCCTG
GTGGACGTGATGCCCTGGCTGCAGTACTTCCCCAACCCGGTGCACCGCTTTTCCGCAATTTCGAGCAGCTCAACCGCAACTTC
15 AGCAACTTCATCTGGACAAGTTCTTGAGGCACTGCGAAAGCCTTCGGCCCCGGGCGCGCCCCCGGACATGATGGACGCTTT
ATCCTCTCTCGGAAAAGAAGGCGCGCGGGACTCGCACGGTGGTGGCGCGCGCTGGATTGGAGAACGTACCGGCCACTATC
ACTGACATCTTCGGCGCAGCCAGGACACCCCTGTCCACCGCGCTGCAGTGGCTGCTCCTCCTCTTACCAGGTATCTGATGTG
CAGACTCGAGTGCAGGCAGAAATGGATCAGGTCTGGGAGGACCGCTTGCCCTGTATGGGTGACCAGCCCAACCTGCCCTAT
GTCTGGCCTTCTTTATGAAGCCATGCGCTTCTCCAGCTTTGTGCCGTGCTACTATTCTCATGCCACCACTGCCAACACCTCT
20 GTCTTGGGTACCAATTTCCCAAGGACACTGTGGTTTGTCAACCAGTGGTCTGTGAATCATGACCCAGTGAAGTGGCCTAAC
CCGAGAACTTTGATCCAGCTCGATTCTTGGACAAGGATGGCCTCATCAACAAGGACCTGACCAGCAGAGTGATGATTTTTC
GTGGGCAAAAGGCGGTGCATTGGCGAAGAACTTTCTAAGATGCAGCTTTTCTCTTCATCTCCATCTCGGCTCACCAGTGGCAT
TTCAGGGCCAACCCAAATGAGCCTGCGAAAATGAATTCAGTTATGGTCTAACCATTAAACCAAGTCATTTAAAGTCAATGTC
ACTCTCAGAGAGTCCATGGAGCTCCTTGATAGTGTCTGCCAAATTTACAAGCCAAGGAACTTGCCAAATAAGAACGAGGCG
25 AAGCTGAAATTTAGAAATATTCACATCTTCGGAGATGAGGAGTAAATTCAGTTTTCAGTTTTCAGTTCTCTTTTGTGCTGCTTC
TCAATTAGCGTTTAAAGTGAGCATAAATCAACTGTCCATCAGGTGAGGTGCTCCATACCCAGCGGTTCTTCATGATGATGG
GCTATGCGAGAGCTTCTGGGAGATTTTGTGAGTCAAGACCTTAAAGGGCCCAATGAATTATTATATACATAGTCAATCTTGGT
TATTTCTGAAGGTAGCATTTCTTTGGAGTTAAATGCACATATAGACACATACACCCAAACACTTACACCAAACTACTGAATGAA
GAAGTATTTTGGTAACAGGCCATTTTGGTGGGAATCCAAGATGGTCTCCCATATGCAGAAATAGACAAAAGTATATTA
30 CAAAGTTTCAGAGTATATTTGTTGAAGAGACAGACAAAGTAATTCAGTGTAAAGTGTGTGATTGAAGGTGATAAGGAAAAGA
TAAAGACCAGAAATTCCTTTTACCTTTTTCAGGAAAATAACTTAGACTCTAGTATTTATGGGTGGATTATCTTTTGCCTTC
TGGTATACTTCTTACTTTTAAAGATAAATCATAAAGTCAGTTGCTCAAAAAGAAATCAATAGTTGAATTAGTGAATATAGTGG
GGTCCATGAGTTATCATGAATTTTAAAGTATGCATTATTAATTTGTAAGTCAAGGTGATGTTGTACCTCTTTTGTCTGCC
AAAGTACAGAAATTTGAATTATCAGCAAGAAAAAAGCCAGCCAAGCTTAAATTTATGTGACCATAATGTACTGATTTTC
35 AGTAAGTCTCATAGGTTAAAAAAGTCAACCAATAGTGTGAAATATATTACTTAACTGTCCGTAAAGCAGTATATAGTATT
ATCTGTTCAGGAAAAGGTGAATAATATATGCTTGTGTAATTTGAAAATTTGAAAAGTACAACTAACGCAACCAAGTGTGCT
AAAAATGAGCTTGATTAATCAACCACCTATTTTTCAGATGGAATGAAGCAGGGTTCTTTTCTTCACTCAAAATTTTGGCGAA
TCTCAAAATTAGATCCTAAGATGTGTTCTTATTTTATAACATCTTATTTGAATTTCTATTTATAATACAGAACTTGTGTTTGA
AAATAACCTAATTAATATATTAATTTCCAAATTCATGGCATGCTTAAATTTTAACTAAATTTTAAAGCCATTTCTGATTATTGA
40 GTTCCAGTTGAAGTTAGTGGAAATCTGAACATTTCTCTGTGGAAGGCAGAGAAATCTAAGCTGTGCTGCCCCAATGAATAATGG
AAAATGCCATGAATTACCTGGATGTTCTTTTACGAGGTGACAAGAGTTGGGACAGAACTCCCATTAACAATGACCAAGTTTC
TCTTCTAGATGATTTTGTGAAAGTTAACATTAATGCTGCTTTTGTGAAAGTCAGAATCAGAAGATAGTCTTGAAGCTGTTTG
GAAAAGACAGTGGAGATGAGGTGAGTTGTTTAAAGATGGCAATTAATTTGTTAGTGGGAAAGCATAAAGCTCAAAATGAA
ATGTATGCATTCATTTTGAAGAGTGAATTTGAAGTTTCAAGTTTAAAGTTTCAATTTGAATTTAACTTCCAAAGAAAGTTCTAC
45 AGTGTCTTAAGTGCTAAGTGCTTATTACATTTTATTAAGCTTTTGTGAATCTTTGTACCAAAATTTTAAAAAAGGGAGTTTGT
ATAGTTGTGTATGTGTGTGTTGGGTGGGGGATGGTAAGAGAAAAGAGAGAAACACTGAAAAGAAGGAAAGATGGTTAAACA
TTTTCCCACTCATTTCTGAATTAATTAATTTGGAGCACAAATTTCAAAGCATGGACATTTAGAAGAAAGATGTTTGGCGTAGCAG
AGTTAAATCTCAATAGGCTATTAATAAGTCTACAACATAGCAGATCTGTTTGTGGTTTGGAAATTTAAAAAATTCATGTA
ATTTTATTTTAAAAATTCATAGCTGTACTTCTGAATATAAAAAATCATGCCAGTATTTTAAAGGCATTAGAGTCAACTACAC
50 AAAGCAGGCTTGCCAGTACATTTAAATTTTGGCACTTGCCATTTCCAAATATTTATGCCCCCAAGGCTGAGACAGTGAAT
TTGGGCTGCTGTAGCCTATTTTATAGATGAGAAATGTGTAGCTGCAAAAATAATCATGAACCAATCTGGATGCCCTATTATG

TCAACCAGGTCCAGATGTGCTATAATCTGTTTTTACGTATGTAGGCCAGTCGTCATCAGATGCTTGCGGCAAAAGAAAGCTGT
GTTTATATGGAAGAAAGTAAGGTGCTTGGAGTTTACCTGGCTTATTTAATATGCTTATAACCTAGTTAAAGAAAGGAAAGAAA
ACAAAAACGAATGAAAAATACTGAATTTGGAGGCTGGAGTAATCAGATTACTGCTTTAATCAGAAACCTCATGTGTTTCTA
CCGGAGAGAGAATGTATTTGCTGACAACCATTAAGTCAGAAGTTTACTCCAGGTTATTGCAATAAAGTATAATGTTTATTA
5 ATGCTTCATTTGTATGTCAAAGCTTTGACTCTATAAGCAAATTGCTTTTTTCCAAAACAAAAGATGCTCAGGTTTGT
GAATTTTCTAAAAGCTTTCATGTCCCAGAACTTAGCCTTTACCTGTGAAGTGTTACTACAGCCTTAATATTTCTCTAGTAGATC
TATATTAGATCAAAATAGTTGCATAGCAGTATATGTTAATTTGTGTGTTTTTAGCTGTGACAACTGTGTGATTAAGGTATA
CTTTAGTAGACATTTATAACTCAAGGATACCTTCTTATTTAATCTTTTCTTATTTTGTACTTTATCATGAATGCTTTTAGTGT
GTGCATAATAGCTACAGTGCATAGTTGTAGACAAAGTACATTCTGGGAAACAACATTTATATGTAGCCTTTACTGTTTGATAT
10 ACCAAATTAAAAAAAATTTGATCTCATTACTTATACTGGGACACCATTACCAAAATAAAAAATCATTTCATAATCTTGAA
AAAA
327
MTTTLVSATIFDLSEVLCKGNKMLNYSAPSAGGCLLDRKAVGTPAGGGFPRRHVSITLPSKHFQNLSSLKGEPAALSSRDS
RFRDRSFSEGGERLLPTQKQPGGQVNSSRYKTELCPFEENGACKYGDCKQFAHGIHELRLTRHPKYKTELCTFHTTIGFCP
15 YGPRCHFIHNAEERRALAGARDLSADRPRLQHSFSFAGFPASAAATAATGLDSDPTITPPPILSADDLLGSPTLPDGTNNPFA
PSSQELASLFAPSMGLPGGGSPTTFLFRPMSESHPMFSDPPSPQDLSLSDQEGYLSSSSSSHSGSDSPTLDNSRRLPIFSRLSIS
DD
328
CGGAGTCAGAAAGGCGAGGGCGCCGGGAACCTGGCGTGTGGGACTCCAGACAGGAGAGGCTGCGCCTTCCCCGCACCGGGACCT
20 TCGCGACACACCAGATCCTCGCCCCCTGGCTCGCGCGAACGCACAGGATGACCACCACCTCGTGTCTGCCACCATCTTCGACTT
GAGCGAAGTTTATGCAAGGGTAACAAGATGCTCAACTATAGTGCTCCCACTGTCAGGGGGTTGCTGTGGACAGAAAGCAGT
GGGCACCCCTGCTGGTGGGGGCTTCCCTCGGAGGCACTCAGTCACCCCTGCCAGCTCAAGTTCCACCAGAACAGCTCCTCAG
CAGCCTCAAGGGTGAGCCAGCCCCGCTCTGAGCTCGCGAGACAGCCGCTTCCGAGACCGCTCCTTCTCGGAAGGGGGCGAGCG
GCTGCTGCCACCCAGAAGCAGCCCGGGGGCGCCAGGTCAACTCCAGCCGCTACAAGACGGAGCTGTGCCGCCCTTTGAGGA
25 AAACGGTGCTGTAAAGTACGGGACAAGTGCCAGTTTCGCACACGGCATCCACAGCTCCGCAGCCTGACCCGCCACCCCAAGTA
CAAGACGGAGCTGTCCCGCACCTTCCACACCATCGCTTTTGCCCTACGGGCCCGCTGCCACTTCATCCACAACGCTGAAGA
GCGCGTGCTGCCCTGGCCGGGCGCCGGGACCTTCTCCGTGACCGTCCCGCTCCAGCATAGCTTTAGCTTTGCTGGGTTTCCAG
TGCCGCTGCCACCGCCGCTGCCACCGGGCTGCTGGACAGCCCATCCATCACCACCCCTATTCTGAGCGCCGATGACCT
CCTGGGCTCACCTACCCTGCCGATGGCACCAATAACCTTTTGCTTCTCCAGCCAGGAGCTGGCAAGCTCTTTGCCCTAG
30 CATGGGGCTGCCGGGGTGGCTCCCCGACCACCTTCTCTTCCGGCCCATGTCCGAGTCCCTTCACATGTTTGACTCTCCCCC
CAGCCCTCAGGATTTCTCTCTCGGACCAGGAGGGTACCTGAGCAGCTCCAGCAGCAGCCACAGTGGCTCAGACTCCCCGACCTT
GGACAACCTCAAGACGCTTCCCATCTTCAGCAGACTTTCCATCTCAGATGACTAAGCCAGGGTAGGGAGGACCTCCTGCCTAC
TCCAGCCCTACCTGACCCACATCCCATACCTCTTCTCCCTACCCATCCCATTTCCCAAGGCCCTACATTAACAAGGTAA
AGCTCAACCCCTTTCCCCCAGCAGCTCAGAAATGTGCCCTCCCTCTCCCCCTCATAACCCACCTAACATAAGGACAAGTCAATT
35 GTCAGTAGCTTCTTCTGGCTTGAACCCCTCCCTGGATTTTATAGCCCACTTACCATGCATAACAGACAAGTCCCATATTTT
GTCAGTAGATGCTTTTCTTCCGGCTTAAAGCTTAAAGTGCCAAATCACAAGAGAAAAGCAGTAACAGTTTACAGAAGCAAC
TTAGTGCTTGTAACTAATCTAATTTGTCACTGTGACTACATTACCTCTTCAGCGCCAGAGGGCACCCGTGGGCCTCCGGAGCCTC
TGCCCATGGCGGGTGGAGACCCGGAACAGCAGCCCCCTCCACTGGCGACACAACCTGCACCTTCCCTCATTTAGTCTCCCGC
ACACTTATTCCTCCTCCCCCTTCCCGGTGGCACCTTCCACCTGTACCCGCCCCCCCCACCCCTGCCCTTGGAA
40 GAGTTGTGCCAGACCAGGTTTGGGGGAAACCTGCTTTGACATTTCAAACCTTTTCTTCCCGATCTGAACCCCTGTTGACT
AATCTTGCTGGGTTTGTGTAGGTCTGCAGGAAGGAAGGCTGAAAAAGCGGACGAAGATTTGACTTAAGTGGGACTTTGTGAT
TTAATTTTTCTTTTTTTTAAAGTGGGGAGGAAGGGAAGCTAGATGGACTAGGAGAGACTTGATTTTGGTGCTAAAGTTCCCA
GTTTATATGTGACATCTTTTTTAAAAAAAATAACAACAAAAAAGCTAGAGAGAAAGCTAAAAAAGTAAGGGGTGAGC
AGTTAATGTGATTTTATCCACATACAATATCTGTGTAACAGATTTCTCTGTAGAAGTAGCTTTAATGGTTTGTGCTCTAGAATA
45 CCGTAGGTCTATCCTTAGAGCACTCAGCCATGCTTTCTTCCCTGGGTTTAAACTTCATATAACTTTTCAGAAATGGAGAGCA
AAAAATTTGCTTGTCACTGCACATCAATATAAAAAAGCTTATTTAACTTATCAAAACGTATTTATGCAAACTATGCTTTTTT
TTGTTAATTTTGTTCATATTTATCGGGATGACAAATCCATAGAATATATTTCTTTATGTTAAATATGATCTTCATATTAATCT
TAAAAATTTGTGACGTGCTTTTTCTTTTTTCCACAGTTTTAATATATTTCTTTCAACGACATTTTTTGTAACTTTTACACT
TTTTTGGTTATTTTATTTTAAAAAATGAAAAATTAATTTAAAAAATGCAAAAGCTGTTGGATTTATTTTAAAGAAATTC
50 CCCCTTTGTGTTGGACTGCAAAATGAGTTTCTTTCTTTTAGGCCCTTCACAACAGGACTGAGAATGTATGTAAAAGTTCTGT

- GACAGTACAGAAGGAAAAACAACCTTTTATGTATAGCTTCTAAAAGGGGAAAAAAGAGAAACCTTTGACTTCCACGT
GCCCATCTCAAGACATTCCACTCACAGATTTGAGGTTCTGGATTCCAGGTCTGGAGTTTCCCAATGTTAATGTAAACAGAACTG
GCACACACACATTAAGATGAATGTAATTATTATTCCTCTTGCTGGTCACTACCGTCGCTTTCTATTTCTCTTTCTTTGTGTGAA
TTTATTTAAAAGAAAAAAACCTTTTGTAAACGACTATTTGCAGTTTAAAAATCAATAAACCCCGTTTTCAGAAACATT
- 5 329
MATFPPATSAPQQPPGPEDESSLDESPLYSLAHSYLGGGGRKGRTKREAAANTNRPSPGGHERKLVTKLQNSERKKRGARR
330
- CAGCCTTATAAGCTGAGGGAGTGGAGAGGCCCGGGGCCAGGAAAGCAGAGACAGACAAAGCGTTAGGAGAAGAAGAGAGGCAGG
GAAGACAAGCCAGGCACGATGGCCACCTTCCCACCAGCAACCAGCGCCCCCAGCAGCCCCAGGCCCGGAGGACGAGGACTCC
10 AGCCTGGATGAATCTGACCTCTATAGCCTGGCCCATCTCTACCTCGGAGGTGGAGGCCGGAAGGTGCGACCAAGAGAGAAGCT
GCTGCCAACACCAACCGCCCCAGCCTTGGCGGCACGAGAGGAACTGGTGACCAAGCTGCAGAATTACAGAGGAAGAAGCGA
GGGGCACGGCGCTGAGACAGAGCTGGAGATGAGGCCAGACCTAGGACTACACCCAGCAATAGAGACGGGACTGCGGAGGAAG
GAGGACCCAGGACAGGATCCAGGCCGGCTTGCCACACCCCCACCCCTAGGACTTATTCGCGTGACTGAGTCTCTGAGGGGCT
ACCAGGAAAGCGCTCCAACCTAGCAAAAGTCAAGATGGGGAGTGGAGGCTGGGAATGGAGGGGCAGGCCAGGAAGATCC
15 CCCAGAAAGAAAGCTACAGAAGAACTGGGGCTCTCCAGGGTGGCAGCAACAATAAATAGACACGCACGGCAGCCACAGCTT
GGGTGTGTGTTTCATCTTGTAAAAAAGGCTTAAAAAAGGCTTAAAAAAGGCTTAAAAAAGGCTTAAAAAAGGCTTAAAAAAGGCTT
331
- MEGISIYTSNDYTEEMSGDYDSMKPECFREENANFNKIFLPTIYSIIFLTGIVGNGLVILVMGYQKKILRSMTDKYRLHLSVAD
LLFVITLPLFWAVDAVANWYFGNLFCKAVHVIYTVNLYSSVLILAFISLDRYLAIHVATNSQRPRKLLAEKVYVGVWIPALLLT
20 IPDFIFANVSEADDRYICDRFYPNDLWVVFQFQHIMVGLILPGIVILSCYCIISKLSHSKGHQKRKALKTTVILILAFFACW
LPYYIGISIDSFILLEIKQGEFENTVHKWISITEALAFFHCCLNPILYAFLGAKFKTSAQHALTSVSRGSSSLKILSKGRKG
HSSVSTESESSSFHSS
332
- GTTTGTGGCTGCGGCAGCAGGTAGCAAAGTGACGCCGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTGGAGAACCAGCGGT
25 TACCATGGAGGGGATCAGTATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGGAACC
CTGTTTCCGTGAAGAAATGCTAATTTCAATAAAATCTTCTGCCACCATCTACTCCATCATCTTCTTAACTGGCATTGTGGG
CAATGGATTGGTCATCTGGTCATGGGTACAGAGAAGAACTGAGAAGCATGACGGACAAGTACAGGCTGCACCTGTCACTGGC
CGACCTCCTCTTTGTTCATCAGCTTCCCTTCTGGGCAGTTGATGCCGTGGCAAAGTGGTACTTTGGGAAGTTCCTATGCAAGGC
AGTCCATGTCTATACAGTCAACCTCTACAGCAGTGCTCTATCTGGCCTTCATCAGTCTGGACCGCTACCTGGCCATCGT
30 CCACGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTATGTTGGCGTCTGGATCCCTGCCCTCTGCT
GACTATTTCCGACTTCATCTTTGCCAACGTCACTGAGGCAGATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTG
GGTGGTTGTGTTCCAGTTTCAGCACATCATGGTTGGCCTTATCCTGCCTGGTATGTGTCATCCTGTCTCTATGCAATTATCAT
CTCAAGCTGTCACTCCAAGGCCACCAAGCGCAAGGCCCTCAAGACCACAGTCACTCCTCATCTCTGGCTTTCTTCGCCTG
TTGGCTGCCTTACTACATTGGGATCAGCATCGACTCCTTCATCTCCTGGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACAC
35 TGTGCACAAGTGGATTTCATCACCAGAGGCCCTAGCTTTCTTCCACTGTGTCTGAACCCATCCTCTATGCTTTCTTGGAGC
CAAAATTTAAACCTCTGCCAGCAGCACTCACCTCTGTGAGCAGAGGGTCCAGCCTCAAGATCCTCTCAAAGGAAAGCGAGG
TGGACATTCATCTGTTTCCACTGAGTCTGAGTCTTCAAGTTTCACTCCAGCTAACACAGATGTAAAGACTTTTTTTTATACG
ATAAATAACTTTTTTTAAGTTACACATTTTTCAGATATAAAGACTGACCAATATTGTACAGTTTATTGCTTGTGAGTTT
TTGTCTTGTGTTCTTTAGTTTGTGAAAGTTAATTGACTTATTTATATAAAATTTTTTTGTTTCATATTGATGTGTGCTAG
40 GCAGGACCTGTGGCCAAGTTCTTAGTTGCTGTATGCTCTGTTAGGACTGTAGAAAAGGGAAGTGAACATTCCAGAGCGTGTAT
GTGAATACGTAAGCTAGAAATGATCCCCAGCTGTTTATGCTAGATAATCTCTCCATTCCTGGAACGTTTTTCTCTGTTCT
TAAGACGTGATTTGTCTGTAGAAGATGGCACTTATAACCAAGGCCAAAGTGGTATAGAAATGTGGTTTTTCAGTTTTTCAGGA
GTGGGTTGATTTCAGCACCTACAGTGTACAGTCTTGTATTAAGTTGTTAATAAAGTACATGTTAACTTACTTAGTGTATG
333
- 45 MALKERIGWRYSLFVGLLQNLIVVFGALLRPIIRGPASPKIVIQENRKEAQYMLENEKTRTSIDSIDSGVELTTSKPNVPTH
TNPELEPKADLQQLVKTSRPRSKKKAPLLDFSILKEKSFICYALFGLLATL
334

- AGCAATTCATGGCTCTGAAGGAGCGCATTGGCTGGAGATACAGCCTCCTCTTCGTGGGCCTACTACAGTTAAACATTGTCGTCT
TCGGAGCACTGCTCAGACCCATCATCATCAGAGGACCAGCGTCACCAAAAATAGTCATCCAGGAAAATCGGAAAGAAGCACAGT
ATATGCTTGAAAAATGAGAAAAACGAACCTCAATAGACTCCATTGACTCAGGAGTAGAACTAACTACCTCACCTAAAAATGTGC
CTACTCACACTAACCCAGAACTGGAGCCGAAGGCAGACCTGCAGCAGGTCCTGGTGAAGACCAGCCCCAGGCCAAGCAAAAAGA
5 AAGCCCCGCTATTAGACTTCTCCATTTTGAAGAGAAAAAGTTTATTTGTTATGCATTATTTGGTCTCCTTGCGACACTGTGAT
TCTTTGCACCTTCTTGTACATCATTCCTCTGGGCATTAGTCTGGGCATTGACCAGGACAGCGCTGCTTTTTTATTATCTACAA
TGGCCATTGCAGAGTTTTCAGGAGGATCGGAGCTGGTTTGTCTCAACAGAGAGCCCATTCGTGTGATTTACATTGAGCTCA
TCTGCGTCATCTTATTGACTGTGTCTCTGTTTGCCTTACTTTTGTCTAGGGAATTCTGGGGTCTAATGTTCATGTAGCGTATTTT
TTGGGTTTATGGTTGGAACAACAGGAGGGACCCACATTCCACTGCTTGTCTGAGGATGATGTTGTGGGCATTGAGAAGATGCTTT
10 CTGCAGCTGGGGTCTACATCTTCATTTCAGAGCATAGCAGGACTGGCTGGACCACCCCTTGCCTGTTGTTGGTGGACCAAGTAA
GACTACTCAGCAGGCGCTTCTACTCTGCGCAGCTGGCATGGCCCTGGCTGCTGTGTGCTCGCCCTGGTGAGACCGTGTAAAGAT
GGGACTGTGCCAGCATCATCACTCAGGTGAAACAAAGGTAGTTCAGCCATCGTGGGAAGACTTTACAGGACATACCTGAAGACTT
TCTGGAATGGATCTTGCAAAAAATGAGCACAGAGTTTCAGCTGCAAAATGGAGCCGGTATGACACACTTTCTTACACCAACAGCC
ACTGTGTTGGCTGGAGAGGGATGGGGTGGGCCAACGGAGACACAAGGAGGTAGAGGAGCTAACCCCTTACTCCACTTTCAAAA
15 ACTACATTTTAAAGGGAATGTGTATGTGAAGAGCACTACCAACATCGCTTTTGTTTTAAAGTTTTCCTTTTGTCTGTGTTTTTAA
GCCAAAACAAAAACAACCAAGCACTCTTCCATATATAAATCTGGCTGTATTTCAGTAGCAATACAAAAGATATGTAGAAAGACT
CTTTGGTTTCACATTCCAATATTAAATAGTGACACGAACGGCAAGTGGTTTAAAGCTTTTACATGGGATAAATGATTTTC
TTTCTTTCTTTTCTTTCTTCGTATGGTCTTGTGAGAATAAACTACTATCTTGAATAAAACAACATCCAACCCAGGTCATTGAAA
TGAAATTGGCCAGTC
- 20 335
- MTVSVVPQREPLVLGGRLAPLGFSSRGYFGALPMVTTAPPLPRI PDPRALPPTLFLPHFLGGDGPCLTPQPRAPAALPNRSLA
VAGGTPRAAPKRRKKKVRASPAQQLPSRFHQYQQRPSLEGGRSPATGPSGAQEVPGPAAALAPSPAAAAGTEGASPDPLAPLR
PAAPGQTPLRKEVLKSKMGKSEKIALPHGQLVHGIHLYEQPKINRQKSKYNLPLTKITSAKRNENNFWDQSVSSDRIQKQEKKP
FKNTENIKNSHLKKS AFLTEVSQKENYAGAKFSDPPSPSVLPKPPSHWMGSTVENSQNRELMVHLKTLKLVQT
- 25 336
- TGTTCCGCGATCTTCTCAGGCTCTCCTAGCAGCATCCATCGCCGCCACCCTATCTTCACTGGCTTCACCTTCTCCTCTCTCTTT
CGTTGCTGAGCGACAAGCTTTCCTAGCGCTATGACTGTCTGCTCCGTCCCGCAGCGGGAGCCGCTCGTCTCGGGTGGCCGCTTG
CGCCGCTTGGCTTTTCCCTCCCGAGGTTACTTTGGGGCCCTCCCGATGGTGACCACGGCTCCGCCTCCTTTACCCCGGATCCCGG
ACCCCGGGCAGTGGCCCCGACCTCTTCTCTCCCTCATTTTCTAGGGGAGATGGCCCGTGTCTGACCCCCAGCCTCGCGCTC
30 CAGCAGCTCTGCCAACCGCAGCCTCGCCGTGGCGGGAGGCACTCTCGGGCAGCGCGAAGAAGCGGCGAAAGAAGAGGTGC
GGGCCAGCCCCGAGGGCAGCTGCCAGCCGCTTCCACCAGTACCAGCAGCACCAGCGGAGTCTGGAGGGCGGGCGGAGCCCCG
CGACCGGCCGAGCGGAGCGCAGGAGTCCCGGGCCCGGCGCCGCTTGGCCCCGAGTCTGCAGCCGACGGCGACGGAGG
GAGCCAGCCCCGACCTTGGCCCCGTGCGGCCCGGCTCCCGGCCAAACCCCCCTCAGGAAAGAGGTTTAAAAATCAAAGATGG
GAAAATCGAGAGAAAATTGCCCTTCCCCATGGCCAGCTTGTTCATGTTATACACTTGTATGAGCAACCAAGATAACAGACAGA
35 AAAGCAAATATAACTTGCCACTAACCAAGATCACCTCTGCAAAAAGAAATGAAAACAACITTTGGCAGGATTCTGTTTCATCTG
ACAGAATTGAGAAGCAGGAAAAAAGCCTTTAAAAATACCGAGAACATTAATAATTCGCATTTGAAAGAAATCAGCATTCTAA
CTGAAGTGAGCCAAAAGGAAAATATGCTGGGGCAAAGTTAGTGATCCACCTTCTCCTAGTGTCTTCCAAAGCCTCCTAGTC
ACTGGATGGGAAGCACTGTTGAAAATTCCAACAAAACAGGGAGCTGATGGCAGTACACTTAAAAACCTCTCAAAGTTCAAA
CTTAGATTTCAGATTTCAGTATGTGTGTAACATAATTTTCCCATATCCCTGGACTCTTGAGAAAATTGGTACAGAAATGGA
40 AATTGCTTGTGCAACATACAAATGCAAAAGATGAGTTTAAAAAATTACATACAAACAGCTTGTATATATTTATATTTTG
TAAATACTGTATACCATGTATATGTGTATATTTGTTCACTTGGAGGATATATATAGTTTGTATGAAAGTATGTATTTTG
CCCTGCCCATATGCAGGTGTTTTTGTATATATAAATGGATAAATTTAAGTGTGTGCTAAGGCACATGGAAGACCGATTATAT
TTGCACAAGGTACTGAGATTTTTTTCAAGAAAACAGCTGTCAAATCTCAAGGTGAAGATCTAAATGTGAACAGTTTACTAATGCA
CTACTGAAGTTTAAATCTGTGGCACAATCAATGTAAAGCATGGGGTTTGTCTTCTTAAATTGATTGTAAATCTGAAATTACTGAA
45 CAACTCCTATTCCCATTTTGTCTAAACTCAATTTCTGGTTTGTGTATATATCCATTCCAGCTTAATGCCTCTAATTTTAAATGCC
AACAAAATGGTTGTAATCAAATTTTAAATAATAATAATTTGGCCCCCTTTTAAATAAGTCTTGACTCTTTGTGTGTGACT
GTTTCTCATGTTTGAATGTGTGACTAGGAGATGATTTTGTGTGGTTGGATTTTTTGTACTTCTACTTTACTGGCTGAGTGTGAG
CCGCCATGCCTGGCCATAATCTACATTTTCTTACCAGGAGCAGCATTGAGGTTTTTGTAGCATAGTACTTGACTACTCTAGAGGC

TGAGACGGGAGCATCTCTTGAGCCTGAGAAGTGAGATTGCAATTGAGCTAGGATCAGGCCACTGCACTCCAGCCTGGGTAACA
GACGCTGTCTCAAAAAAAGGCCAAGAGAAAGTAAGGAGACAGA

337

MAEVEDQAARDMKRLEEKDKERKNVKGIRDDIEEDDQEAIFYRYMAENPTAGVVQEEEEEDNLEYDSDGNPIAPTCKIIDPLPPI
5 DHSEIDYPPFEKNFYNEHEEITNLTQQQLIDLRLHKLNLRVSGAAPPRPGSSFAHFGFDEQLMHQIRKSEYQTPTIQCCQGVVPA
LSGRDMIGIAKTGSGKTAAFIWPMLIHIMDQKELEPGDGPIAVIVCPTRELCCQIHAECKRFGKAYNLSVAVYGGGSMWEQAK
ALQEGAEIVVCTPGRLLIDHVKKKATNLQRVSYLVFDEADRMFDMGFYQVRSIASHVRPDRQTLFSATFRKKIEKLARDILID
PIRVVQGDIGEANEDVTQIVEILHSGPSKWNWLTRRLVEFTSSGSVLLFVTKKANAEELANNLKQEGHNLGLLHGDMDQSERNK
VISDFKKKDI PVLVATDVAARGLDI PSIKTVINYDVARDIDTHTHRIGRTGRAGEKGVAYTLLTPKDSNFAGDLVRNLEGANQH
10 VSKELLDLMAQNAWFRKSRFKGGKGLNIGGGGLGYRERPGLGSENMDRGNNVMSNYEAYKPSTGAMGDRLTAMKAAFQSQY
KSHFVAASLSNQKAGSSAAGASGWTSGASLNSVPTNSAQQGHNSPDSPTSAKGI PGFGNTGNI SGAPVTPSAGAQGVNNTA
SGMNSREGTGGSGNKRERYTENRGSSRSHGETGNRHS DSPRHGDGGRHGDGYRHPRESSRHTDGRHGENRHGGSAGRHHENR
GANDGRNGESRKEAFNRESKMEPKMEPKVDSSKMDKVDSDTKTADGFAVPEPPKRRKSRWDS

338

15 GTCCATCTTGCTGCTGTCCACTTTGGGTTCCATCTTGGGCTCCATCTTGCTCTCACGATTAAAAGCTTCTTTCTGCTTTCCCC
ATTCCGACCATCATTTGCACCCCGGTTCTCCCATGCGGGCTGCGCTTCTCCATGTCTGTCTCCCCGTGCCGATGGCCATC
AGTATGACGGCTGCTGTCTTCTGGATGGCGGTATCCATCTCCATGGCGACCACTATCCGTGACGTGGACTATCGCTATGCCG
ATTGCCAGTCTCTCCGTGACTGTGACGGCTGTGCCCCGGTTCTCAGTATATCTCTCTCTTTTCCCGTGTGCTCCCCCAGTCCC
TTCTCGGCTGTTATTCCCTGAAGCTGTGTGTTGACTCCTTGGGCTCCGGCAGACGGGTAGGTACACAGGGGACCACTGATGTT
20 GCCAGTATTGCCAAAGCCTGGGATGCCCTTGGCGGCACTGGTGACGGGGCTGTGAGGACTGTTATGGCCCTGTTGTGCTGAGTT
AGTTGGAACAGAATTCAAGCTCCCTGCACTAGTCCACCTGCCCACTGCTTCTTCGCTTCTCTCTTGGAAAGTCCAGTCTCTCC
TCGGCTTGCAATGGACCCCACTGCTCCTGCGCCGCTGGTGCTCTCTGCACTGCGCTGGTCTCTGCAAGTGCAAGAGTGCAA
ATGCACCTCTGCAAGAAGAGCTGCTGCTCTGCTGCCCCGTGGGCTGTAGCAAGTGTCAGGGCTGTGTTGCAAGGGGGC
GTCAGAGAAGTGCAGCTGCTGCGACTGATGCCAGGACAACCTTTCTCCAGATGTAAACAGAGAGACATGTACAAACCTGGATT
25 TTTTTTTTATACCACCTTGACCATTGTCTACATTCTTTCTCTGTGAAATATGTGAGTGATAATTAAACACTTTAGACCTGAA
AA
AA

339

MTQTLKYASRVFHRVRWAPELGASLGYREYHSARRSLADIPGPSTPSFLAELFCKGGLSRLHBLQVQGAHFGPVWLASFQTVR
30 TVVVAAPALVEBLLRQEGPRPERCSFSPWTEHRRRCRQRACGLLTAEGEEWQRLRSLAPLLLRPQAAARYAGTLNNVCDLVR
LRRQRGRGTGPPALVRDVAGEFYKFLGELIAAVLLGSRGLGLEAQVPPDTETFI RAVGSVVFVSTLLTMAMPWHLRHLVPGPWGR
LCRDWDQMFQFAQRHVERREAEAAAMRNGQPEKDLESGAHLTHFLFREELPAQSILGNVTELLLAGVDTVSNTLSWALYELSRH
PEVQTALHSEITAAALSPGSSAYPSATVLSQLPLLKAVKEVLRLYPVVPNGNSRVDPKDIHVGDYII PKNTLVTLCHYATSRDPA
QFPEPNSFRPARWLGEGETPHPFASLPFGFGKRSCMRRLAELELQMALAQILTHFEVQPEPGAAPVRPKTRTVLVPERSINLQ
35 FLDR

340

GGGGTTGAGATATGATGCTCAGGAGAAGCGCTTTCTTTTCGCGAGCACCTTGAACCAGACCATGACCCAGACCTCAAGTACGCC
TCCAGAGTGTTCATCGCGTCCGCTGGGCGCCGAGTTGGGCGCCTCCCTAGGCTACCGAGAGTACCACTCAGCACGCCGGAGC
40 TTGGCAGACATCCCAGGCCCTCTACGCCAGCTTTCTGGCCGAAGCTTTCTGCAAGGGGGGGCTGTGAGGCTACACGAGCTG
CAGGTGCAGGGGCCCGCACTTCGGGCGGTGTGGCTAGCCAGCTTTGGGACAGTGCGCACCGTGTACGTGGCTGCCCTGCA
CTCGTCGAGGAGCTGCTGCGACAGGAGGACCCCGCCCGAGCGCTGCACTTCTCGCCCTGGACGGAGACCCGCCGTGCCGC
CAGCGGGCTTGGGACTGCTCACTGCGGAAGGCAAGAAATGGCAAAGGCTCCGAGTCTCTGGCCCCGCTCTCTCCGGCCT
CAAGCGGCCGCCGCTACGCCGAACCTGAACAACGTAGTCTGCGACCTTGTGCGCGCTGTGAGGCGCCAGCGGGGACGTGGC
ACGGGGCCGCCGCCCTGGTTCCGGACGTGGCGGGGAATTTTACAAGTTCGGACTGGAAGGCATCGCCCGGTTCTGCTCGGC
45 TCGCGCTTGGGCTGCTGAGGCTCAAGTGCCACCCGACACGGAGACCTTCATCCGCGCTGTGGGCTCGGTGTTGTGCTCCAG
CTGTTGACCATGGCGATGCCCCACTGGCTGCGCCACCTTGTGCTGGGCTTGGGCGCCCTCTGCGGAGACTGGGACCATG
TTTGCAATTGCTCAGAGGCAGTGGAGCGCGAGAGGCAGAGGCAGCCATGAGGAACGAGGACAGCCGAGAAGGACCTGGAG
TCTGGGGCGCACTGACCACTTCTGTTCGGGAAGAGTTGCTGCGCCAGTCCATCTGGGAAATGTGACAGAGTTGCTATTG

CGGGAGTGGACACGGTGTCCAACACGCTCTCTTGGGCTCTGTATGAGCTCTCCCGGCACCCCGAAGTCCAGACAGCACTCCAC
TCAGAGATCACAGCTGCCCTGAGCCCTGGCTCCAGTGCCTACCCCTCAGCCACTGTTCTGTCCCAGCTGCCCTGTGAAGCG
GTGGTCAAGGAAGTGTAAAGACTGTACCCTGTGGTACCTGGAATTTCTCGTGTCCCAGACAAAGACATTCATGTGGGTGACTAT
ATTATCCCCAAAAATACGCTGGTCACTCTGTGTCACTATGCCACTTCAAGGGACCCTGCCAGTTCCAGAGCCAAATTTCTTTT
5 CGTCCAGCTCGCTGGCTGGGGGAGGGTCCACCCCCACCCATTTGCATCTCTTCCCTTTGGCTTTGGCAAGCGCAGCTGTATG
GGGAGACGCCTGGCAGAGCTTGAATTGCAATGGCTTTGGCCAGATCTTAACACATTTTGAGGTGCAGCCTGAGCCAGGTGCG
GCCCCAGTTAGACCCAAGACCCGACTGTCTGGTACCTGAAAGGAGCATCAACCTACAGTTTTTGACAGATAGTCCCATGGA
AAGAGACTGTATCATCACCTTTTCATTCATCATAGGGATAAGATTTTGTAGGCACAAGACCAAGGTATACATCTTCCCCTA
ATGCCCTATCTGACCAAACTGGATAGAACCCATAGTGAAGTGTGAGGCGGCCCTGACCAATGTGTGAAGTATGCACCTTGGCCT
10 GACTCAGGAAGCCAGGTGAGAAAACCATGGTCTCTCTGCTTGGCTTGGCCCTTCTGATCATGTATGCAATCCCCAAGGATGAAAT
CAGATTTTAACTAATAATGCTGGATGGCTGAGGAAAGATTCAACTGCCTCTCTTTTGGGCTTTTCATAGTGTTCATGATGCT
GCTGGCTAAGCATTTATCAAAGCATAAGCTCAGTAACGTGTGCATCTGGTCTGTACCTGGTTGGTCTTCTGCTTTTGCATGTAAG
CTCTTTGAGAGGAAGGGTGAAGCCTTATTTGTTTTTATGTCCCTGCCAGGGCCTGTCTCTGACTAGGTGTACCATAACAT
TCTTAGATTGAATCTGAACCATGTGGCAGAAGGGATAAGCAGCTTACTTAGTAGGCTCTGTCTACCCCTTCTTCTTTGTCTT
15 GCCCCTAGGAAGGTGAATCTGCCCTAGCCTGGTTTACGCTTTCTTATAACTCTCCTTTGCTCTCTGGCCACTATTAAGTGGGT
TGCCCCATCACTTAGTTCTCAGGCAGAGACATCTTGGGCCGTGCCCTGCCAGGCCTCTGGCTTTTATATGAAAATTTTAA
AATATTCAAAATTTTAGAATAAATCAAATATTCATCTCT

341

MRAHEGREIPSLGGARRREVLQAGRSQRAAGRRRRRQLELELVGSGRPGGPPPGPGRRTCAAALPPEWPRRTGLPRRGPRP
20 PLAMAKWLNKYFSLGNSKTKSPQPDPYREQRRRGERPSQPQAVPQASSAASACGPATASCFSASSGSLPDDSGSTSDLI
RAYRAQKERHFQDPYNGPGSSLRKLAMRLDYCGSGEPGQVQRAFSSASSAGAAGCCASSGAGAAASSSSSGSPHYLRSS
SERRPATPAEVYISPKHRLIKVESAAAGGAGDPLGGACAGGRWSPACGGKLLNKCAASAAEESGAGKDKVTIADDYSDP
FDAKNDLKSAGGESAGYMEPYEAQRIMTEFQREQSVRSQHKGIQLYDTPYEPEGQSVSDSESTVSPRLRESKLPQDDDRPA
DEYDQPEWNRVTSPALAAQFNGNEKRQSSPSPSRDRRQLRAPGGGFKPIKHGSPFCGILGERVDPVPLEKQIWIYHGAISR
25 GDAENLLRLCKECSYLVNSQTSKHDYPLSLRSNQFMHMKLAKTKEKIVLQNSPPFDSVPEVIHYTTRKLPKGAEHLSSL
YPVAVRTL

342

CGGGCCGCGGGACGGGACGGGCGCGGGCTCCGGCGGGCGCGGGCTTCTCCGTCGCTCGCTGTCTCTCCCGGCCG
ATTTCTCTCCGCTGCGGGCCGAGCTCTCCCGAGCTCGCAGGAAGGAAGAAGGAGCCGAGGACGCCGAGAAGTTCCCGCGG
30 CAGCCGCGGATCCCGGCAAGGCGGAGGCTGCGGCTCCGACGGGCGAGGAGCGGATCCACGGCGAGGGCGTACGGCCAAAGG
GTCCGCGCGCTGGAGCGCTCGGACCTTCCGCTCTCCCCGGGCGTGGGCGGGACCCATGAGACGCGCCACGAGGGCGCGA
GATTCCTAGCTTGGGCGGCGCTAGGCGGAGGAGGTGTGTCAGGCGGCGGAGCCAGAGAGCTGCCGCGAGGAGCGCGCGG
GCAAGAACTTGAACCTGGCGTGGGAGCGGGCGCCCCGAGGCCCCCGCGGGCGCGGGCGCGAGGACCTGCGCGCAGC
GCTGCCCCCGAATGGCCCGCGCGCGGACCGGGCTCCCGCGCGCGGCGCTAGGCGGCTCTCGCCATGGCCAAGTGCTAAA
35 CAAGTACTTCAGCTTGGGCAACAGCAAGACCAAGAGCCCCCGAGCCCGCGGCGGCACTACCGCGAGCAGCGGCGCGGAGG
CGAGCGGCTTCGAGCCCCCCCCAGGCGGTGCGCGAGGCTCTCCCGCGCTCGGCGTCTGCGGTCCGGCCACCGCTCTCTG
CTTCTCAGCTCTTCGGGCTCGCTGCCCCGACGACAGCGGCGAGCACCAGCAGCTCATCCGCGCTACCGCGCGCAGAAGGAGCG
ACACTTCCAGGACCCCTACAACGGGCTGGCTCGTCGCTGCGCAAACTGCGCGCATGTGCGGCTGGACTACTGCGGCGGAG
CGGGGAGCCAGGCGGGTCCAGCGCGCTTCTCGGCTCGTCCGCGTGGGCGCGCGGGCTGTGTGCTGCGCTCTCTGGGCGG
40 GGGCGCGCGCGCTCTCGTCTCTCGTCTCCGCTCTCCGATCTCTACCGAGCAGCAGCGAGCGGCGGCGGCCACGCGCGG
CGAGGTGCGCTACATCTCCCCAAGCACCCTCATCAAGTGGAGAGCGCGCGGCGGTGGGCGCGGGACCCCTGGGGG
CGCTGCGCGGCGCGCGCACCTGGAGCCGAGCGGCTGCGGAGGCAAGAACTGCTCAACAAGTGGCGCGCTCAGCGCGGGA
GGAGAGCGGGCGCGCAAGAAAGGACAAGGTGACCATAGCCATGACTACTCAGATCCCTTTGATGCCAAGAATGATCTCAAGAG
CAAAGCAGGAAAGGGGAGAGTGCTGGCTACATGAGCCCTATGAGGCACAGAGGATCATGACAGAATTTAGAGGCAGGAAAG
45 TGTCGGTCCCAGCATAAAGGTATCCAGTTATATGACACCCCTTACGAACCTGAAGGCCAAAGTGTGACTCGGACTCGGAGAG
CACAGTCAGCCCCGACTGCGGGAGAGCAAGCTGCCCCAGGATGACGACAGGCGCGGATGAGTACGACACAGCTTGGGAGTG
GAACCGGGTACCAGCCAGCCCTGGCAGCACAGTTTAAATGGCAACGAGAAGCGGCAGTCATCCCCCTCACCTTCCGGGACCG
GCGGCGCAGCTTCTGTCCTTGGAGGGGCTTTAAGCTTATCAACATGGGAGCCCTGAGTTCTGCGGATCTTAGGAGAAAG
GGTGGATCTGCGCTCCCCCTGGAGAAGCAATATGGTATCACGAGCCATCAGCAGAGGAGACCGGAGAACCTGCTGCGACT
50 CTGCAAGGAGTGTAGCTACCTTGTCCGGAACAGCCAGACAGCAAGCATGACTACCCCTCTCCCTGAGGAGCAACAGGGTTT

- TATGCACATGAACTGGCCAAAACCAAAGAGAAATACGTTCTGGGTGAGAACAGCCCTCCGTTTCGACAGTGTCCCGGAAGTCAT
CCACTACTACACCACCAGAAAGCTACCCATCAAAGGGGCTGAGCACTTGTCCCTCCTCTATCCCGTGGCTGTGAGGACCCCTGTG
AGCGGACCAGACCTGCCCTGCTCTGTGACAGAGCCTGGAGACTTGGAGGTGCCAGAGGCCCCCACCACAGCCCCAGCCAC
TGTTCCTGGCTGTGTCTGTTTGTGTGTGTATGGTACTAGCACACCACTGCATGTCTCTAGAATGTGTGTGCCACTTACGGGG
5 GCTGGAGGCCCTGGATAAAGACAGAAGGGCGGCAACACC
- 343
- MASVLSRRLGKRSLLGARVLGPSASEGPSAAPPSEPLLEGAAPOFFTTSDDTPCQEQPKEVLKAPSTSGLQQVAFQPGQKVYVW
YGGQECTGLVEQHSWMEQVTVWLLLEQKLQVCCRVEEVWLAELQGPCPOAPPLEPGAQALAYRPVSRNIDVPKRKSDAVEMDEM
MAAMVLTSLSCSPVVQSPPGTEANFSASRAACDPWKESGDISDSGSSSTTSGHWSGSSGVSTPSPPHQASPKYLGDAFGSPQTD
10 HGFETDPPFLDEPAPRKRKNSVKVMYKCLWPNCGKVLRSIVGIKRHVKALHLGDTVDSQDFKREEDFYFTEVQLKEESAAAA
AAAAAGTPVPGTPTSEPAPTPSMTGLPLSALPPPLHKAQSSGPEHPGESSLP SGALS KSA PGSFWHIQADHAYQALPSFQIPV
SPHIYTSVSWAAAPSAACSLSPVRSLSFSEPPQAPAMKSHLIVTSPPPAQSGARKARGEAKKCRKVYGIHRDQWCTACRW
KKACQRFLD
- 344
- AAGTGCAGCATGTGCGCGAGGAGTCGCTCGGGCACTTATTGAGCGCCGACTGTCTACGGGCGGCCGGGGGTGATGGGCAGAGGCT
TCAGTGTCCCTTCGCTCCGAGGAGAGAGAGGAGCAGCAGCATGGCGAGTGTCTGTCCGACGCCCTGGAAAGCGGTCCCTC
CTGGGAGCCCCGGGTGTGGGACCCAGTGCCTCGGAGGGGCCCTCGGCTGCCCCACCCTCGGAGCCACTGTAGAGGGGCCGCT
CCCCAGCCTTTACACCCTCTGATGACACCCCTGCCAGGAGCAGCCCAAGGAAGTCTTAAGGCTCCAGCACCTCGGGCCTT
CAGCAGGTGGCCTTTACGCTTGGCAGAAAGTTTATGTGTGTGTACGGGGGTCAAGAGTGACAGGACTGGTGGAGCAGCACAGC
20 TGGATGGAGGGTCAAGTGTACCGTCTGGCTGTGAGCAGAAGTGTGAGGTCTGCTGCAGGGTGGAGGAGGTGTGGCTGGCAGAG
CTGCAGGGCCCTGTCCCCAGGACCAACCCCTGGAGCCCGAGCCAGGCCCTGGCCTACAGGCCCGTCTCCAGGAACATCGAT
GTCCCAAGAGGAAGTCGGACGCAGTGGAAATGGATGAGATGATGGCGGCCATGGTGTGACGTCCCTGTCTGCAGCCCTGTT
GTACAGAGTCTCTCCGGGACCGAGGCCAATTTCTGTCTCCCGTGCAGCCCTGCGACCCATGGAAGGAGAGTGGTGACATCTCG
GACAGCGGCAGCAGCACTACCAGCGGTCAGTGGAGTGGGAGCAGTGGTGTCTCCACCCCTCGCCCCCACCAGGCCAGC
25 CCCAAGTATTTGGGGGATGCTTTTGGTTCTCCCCAACTGATCATGGCTTTGAGACCGATCCTGACCTTTCTGCTGGACGAA
CCAGCTCCACGAAAAAGAAAGAACTCTGTGAAGGTGATGTACAAGTGCCTGTGGCCAACTGTGGCAAAGTTCTGCGCTCCATT
GTGGGCATCAAACGACACGTCAAAGCCCTCCATCTGGGGGACACAGTGGACTCTGATCAGTTCAAGCGGAGGAGGATTTCTAC
TACACAGAGGTGCAGCTGAAGGAGGAATCTGCTGTCTGCTGTCTGCTGTCTGCTGCGCAGGCACCCCACTCCCTGGGACTCCACC
TCCGAGCCAGCTCCACCCCCAGCATGACTGGCTGCTCTGTCTGCTCTTCCACCACCTCTGCACAAAGCCAGTCTCCGGC
30 CCAGAACATCTGGCCCGAGTCTCTCCCTGCGCTCAGGGGCTCTCAGCAAGTCAGCTCCTGGGTCTTCTGGCACATTCAGGCA
GATCATGCATACCAGGCTCTGCCATCTTCCAGATCCCACTCTCAACACATCTACACCAGTGTGAGTGGGCTGCTGCCCCC
TCCGCGCCTGCTCTCTCTCTCCGGTCCGGAGCCGCTGCTAAGCTTCAGCGAGCCCCAGCAGCCAGCACCTGCGATGAATCT
CATCAGATCGTCACTTCTCCACCCCGGGCCAGAGTGGTGCAGGAAAGCCGAGGGGAGGTAAGAAGTCCCGCAAGGTGTAT
GGCATCGAGCACCGGACAGTGGTGCACGGCTGCGGTTGAAGAAGCCCTGCCAGCGCTTTCTGGACTGAGCTGTGCTGCAG
35 GTTCTACTCTGTTCTTGGCCCTGCGGCGAGCACTGACAAGAGGCCAGTGTGTACAGCCCTCAGCAGAAACCGAAAGAGAAA
GAACGGAAACACCGAGTTTGGGCTCTGTGGCTAAGGTGTAACTTAAAGCAATTTCTCCATTGTGCGAATTTTATTTT
TTAAAAAAGAAACAAAAATATTTTCCCCCTAAATAGGAGAGAGCCAAAAGTACCAAGGCTATTACAGAGTGAACCAAGT
ACTAAAGAATTAAATACCTCCGTTTCCACATCCCACTCTTAGGGGATTAGCTTGTGCGTGTCAAAGAAGGAACAGCTCG
TTCTGCTCTCTGCTGAGTCCGTGAATTTCTTGTCTTCTAAACTCTTCCAGAAAGGACTGTGAGCAAGATGAATTTACTTTCTT
- 40 345
- MAMHFIFSDTAVLLDFWVSVHSPAGMALSVLVLLLLAVLYEGIKVKAKLLNQVLVNLPTSISQQTIAETDGSAGSDSPVGR
THHRWYLCHFQSLIHVIQVIVIGYFIMLAVMSYNTWIFLGVVLGSAVGYLAYPLLSTA
- 346
- TCGGCACAGGAGCGAGGAGACCCGAGAGCAGACGCGCCCTGGCGCCCGCCCTGCGCAGTCACCATGGCGATGCATTTTCATCTTC
45 TCAGATACAGCGGTGCTTCTGTTTGTATTTCTGGAGTGTCCACAGTCTGCTGGCATGGCCCTTTCCGGTGTGGTGTCTCTGCTT
CTGGCTGTACTGTATGAAGGCATCAAGGTGGCAAAGCCAGCTGTCTCAACCAGGTACTGGTGAACCTGCCAACCTCCATCAGC
CAGCAGACCATCGCAGAGACAGACGGGACTCTGCAGGCTCAGATTCAATCCCTGTTGGCAGAACCCACCAGGTGGTACTTG
TGTCACCTTTGGCCAGTCTCTAATCCATGTATCCAGGTGGTCACTCGGCTACTTCATCATGTGCGCGTAATGTCTACAACACC

TGGATTTTCCTTGGTGTGGTCTTGGGCTCTGCTGTGGGCTACTACCTAGCTTACCCACTTCTCAGCACAGCTTAGATGGTGAGG
AACGTGCAGGCACTGAGGCTGGAGGGACATGGAGCCCCCTCTCCAGACACTATACTTCCAACCTGCCCTTTCTTCTGATGGCTA
TTCTCTCCACCTTATTCCAGCCCCCTGAAACCTTTGAGCTGAAGCCAGCACTTGCTCCCTGGAGTTCCGGAAGCCATTGCAGCAAC
CTTCCTTCTCAGCCAGCCTACGTAGGGCCCAGGCATGGTCTTGTGTCTTAAGACAGCTGCTGTGACCAAAGGGAGAATGGAGAT
5 AACAGGGGTGGCAGGGTFACTGAGCCCATGACAATGCTTCTCTGTGACTCAAACCAGGAATTTCCAAAGATTTCAAGCCAGGGA
GAAGGGTTCCTGGTGTATGCAGGGCATGGAACCTGGACACCCTCAGCTCTCCTGCTTTGTGCTTATCTACAGGAGCATCGCCCCA
TTGGACTTCCTGACCTCTTCTGTCTTTGAGGGACAGAGACCAAGCTAGATCCTTTTCTCACCCTTCTGCTTTGGAACACATG
AAGATCATCTCGTCTATGGATCATGTTGACAACTAAGTTTTTTTTTATTTTTTCCCATTTGAACCTTAGTTGGCAATTTTGCACA
TTCATACAAAAAATTTTTAATGAAATGATTTTCATTGATTTCATGATGGATGGCAGAACTGCTGAGACCTATTTCCCTTTCTTG
10 GGGAGAGAATAAGTGACAGCTGATTAAAGCCAGAGACACAGGACTGCTTTCAGGCTCCTGGTTTATTTCTGATTGACTGAGCT
CCTTCCACCAGCAAGGCACTGCCGTCAGGAAGAAGATGATCTGATGGCCGTGGGTGCTCTGGGAAGCTCTTCGTGGCCCTCAATGCC
CTCCTTTATCCTCATCTTTCTTCTATGCAGAACAAAAAGCTGCATCTAATAATGTTCAATACCTTAATATTTCTCTATTATTACT
TACTGCTTACTCCTAATGATCTAGTGGGAAACATGATTCATTCACTTAAAACTGATTGAAGCCATGGGCAGGTACTGACTGA
AGATGCAATCCAAACCAAGCCATTACATTTTTTGTAGTTAGATGGGACTCTCTGGATAGTTGAACCTCTTCACTTTTATAAAAAAG
15 GAAAGAGAGAAAATCACTGCTGTATACATAAATACCTCACAGATTAGATGAAAAGATGGTTGTAAGCTTTGGGAATTAATAACAA
ATACATTTTAGTAAATAT

347

MSRFVQDLSKAMSQDGASQFQEVIRQLELSVKKELEKILTTASSHEFEHTKKDLDFRKLPHRFLQEKGPSVDWNGKIQRPPED
SIQPYEKIKARGLPDNISSVLNKLVVVKLNGGLGTSMGCKGPKSLIGVRNENTFLDLTVQQIEHLNKTNTDVLPLVLMNSFNTD
20 EDTKKILQKYNHCRVKIYTFNQSRYPINKESLRPAKDVSYSGENTEAWYPPGHGDIYASFYNSGLLDTFIGEGKEYIFVSN
DNLGATVDLYILNHLINPPNGKRCEFMVETNKNTRADVKGGLTQYEGKLRLLVEIAQVPKAHVDEFKSVSKFKIFNTNWLWISL
AAVKRLQBNADIMEIIVNAKTLDGGLNVIQLETAVGAAIKSFENSLGINVPRSRFLPVKTTSDLLLVMSNLYSLNAGSLTMS
KREFPTVPLVKLGSSFTKVQDYLRRFESIPDMLELDHLTVSGDVTFGKNVSLKGTVIIIANHGRIDIPGAVLENKIVSGNLR
ILDH

25 348

AGGAGAGGAAGAGAGACCTGCCCTGTAGCGTGACTCCTCTAGAAAAAAGCCGGAGTATTTTACTAAGCCCTTAA
ATGTCGAGATTTGTACAAGATCTTAGCAAAGCAATGTCTCAAGATGGTGCTTCTCAGTTCCAAGAAGTCATTCCGCAAGAGCTA
GAATTATCTGTGAGAAGGAAGCTAGAAAAATATCTACCCACAGCATCATCACATGAATTTGAGCACACCAAAAAAGACCTGGAT
GGATTTCCGAAGCTATTTTCATAGATTTTGTGAAGAAAAGGGGCTTCTGTGGATTGGGGAAAAATCCAGAGACCCCCGGAAGAT
30 TCGATTCAACCCATGAAAAGATAAAGGCCAGGGGCTGCCGTGATAATATATCTTCCGTGTTGAACAACTAGTGGTGGTGAAA
CTCAATGGTGGTTTGGGAACCAGCATGGGCTGCAAGGCCCTAAAGTCTGATTGGTGTGAGGAATGAGAATACCTTTCTGGAT
CTGACTGTTTCAGCAAATGAACATTTGAACAAAACCTACAATACAGATGTCCCTCTTGTTTTAAATGAACCTTTTAACACGGAT
GAAGATACCAAAAAATACTACAGAAGTACAATCATTGTCGTGTGAAAACTTCACTTTCAATCAAAGCAGGTACCCGAGGATT
AATAAAGAATCTTTACGGCTGTAGCAAAGGACGTGCTTACTCAGGGGAAAAATACAGAAGCTTGGTACCTCCAGGTCATGGT
35 GATATTTACGCCAGTTTCTACAACCTCTGGATTGCTTGATACCTTTATAGGAGAAGGCAAGAGTATATTTTGTGTCTAACATA
GATAATCTGGGTGCCACAGTGGATCTGTATATCTTAATCATCTAATCAACCCACCAATGGAAAAACGCTGTGAATTTGTCTATG
GAAGTCACAAAATAAACACGTGCAGATGTAAAGGGCGGGACACTCACTCAATATGAAGGCAAACTGAGACTGGTGGAAATTTGCT
CAAGTGCCAAAAGCACATGTTGACGAGTTCAAGTCTGTATCAAAGTTCAAAATATTTAATACAAAACCTATGGATTTCTCTT
GCAGCAGTTAAAAGACTGCAGGAGCAAAATGCCATTGACATGGAAATCATTTGTGAATGCAAGACTTTGGATGGAGGCTGAAT
40 GTCATTCAATTAGAACTGCAGTAGGGGCTGCCATCAAAGCTTTGAGAATTTCTTAGGTATTAATGTGCCAAGGAGCCGTTTT
CTGCCGTGCAAAACCACATCAGATCTCTTGTGCTGATGTCAAACCTCTATAGTCTTAATGCAGGATCTCTGCAATGAGTGAA
AAGCGGGAATTTCTTACAGTGCCTTTGGTTAAATTAGGCAGTTCTTTTACGAAGGTTCAAGATTATCTAAGAAGATTTGAAAGT
ATACCAGATATGCTTGAATTTGGATCACCTCACAGTTTCAGGAGATGTGACATTTGAAAAAATGTTTCATTAAAGGGAACCGTT
ATCATCATGCAAAATCATGGTGACAGAAATGATATCCCACCTGGAGCAGTATTAGAGAACAAGATAGTGTCTGGAAACCTTCGC
45 ATCTTGACCACCTGAAATGAAAAATACTGTGGACACTTAAATAATGGGCTAGTTTCTTACAATGAAATGTTCTCTAGGATTTAG
GCACTAAAAGGTACTTTTACTATGTTACTGTACCTGCAGTGTGATTTTAAAAATAGAGTTTCTGCAATATGCTTTTACTGCTA
AGAAAAGCACAGATGGTGCAATACTTTCTTCTTTGAAGAGATCCCAAAGTTAGTTACTCTTAAGTGC

349

5 MKLWVSALLMAWFGVLSVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWANKMEALTSKSAADAEGYLAHPV
NAYKLVKRLNTDWPALDLVLQDSAAGFIANLSVQRQFFPTDEDEIGAALKMLRLQDTYRLDPGTISRGLPGTKYQAMLSVDD
CFGMGRSAYNEGDIYHTVLWMEQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLLSLDPSHERAGGNLRYFEQ
LLEEREKLTNTQTEAELATPEGIYERPDYLPERDVYESLCRGEVVKLTPRRQKRLFCRYHHGNRAPQLLIAPFKEDEWDSF
HIVRYDYVMSDEEIERIKEIAKPKLARATVRDPKTGVLTVASYRVSKSSWLEEDDDPVVARVNRMRQHITGLTVKTAELLQVAN
YGVGGQYEPHDFSRNDERDTFKHLGTGNRVATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGBEGDYRTHAAC
PVLVGCKWVSNKWFHERGQEFRLPCGSTVD

350

10 GGGGAAGGAACACTGTAGGGGATAGCTGTCCACGGACGCTGTCTACAAGACCCTGGAGTGAGATAACGTGCCTGGTACTGTGCC
CTGCATGTGTAAGATGCCAGTTGACCTTCGCAGCAGGAGCCTGGATCAGGCACTTCCTGCCTCAGGTATTGCTGGACAGCCCCA
GACACTTCCCTCTGTGACCATGAACTCTGGGTGTCTGCTGATGGCCTGGCTTGTGCTGCTGAGCTGTGTGACAGCCGA
ATTCTTACCTCTATTGGGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATCCTTGT
GGAGGAAGCCAAGCTTTCAGAGTTAAGAGCTGGGCCAACAAAAATGGAAGCCTTGACTAGCAAGTCAGCTGCTGATGCTGAGGG
CTACCTGGCTCACCTGTGAATGCCTACAACTGGTGAAGCGGCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCTGCA
15 GGACTCAGTGCAGGTTTATCGCCAACCTCTCTGTGCAGCGGCTGCTTCTCCCACTGATGAGGACGAGATAGGAGCTGCCAA
AGCCCTGATGAGACTTCAGGACACATACAGGCTGGAGCCAGGCACAATTTCCAGAGGGGAACCTCCAGGAACCAAGTACCAGGC
AATGCTGAGTGTGGATGACTGCTTGGGATGGGCGCTCGGCTTACAATGAAGGGGACTATTATCATACGGTGTGTGGATGGA
GCAGGTGCTAAAGCAGCTTGATGCCGGGAGGAGGCCACCACAACCAAGTCACAGGTGCTGGACTACCTCAGCTATGCTGTCTT
CCAGTTGGGTGATCTGCACCGTGCCTGGAGCTACCCGCCCTGCTCTCCCTTGACCCAAGCCACGAACGAGCTGGAGGGAA
20 TCTCGGTACTTTGAGCAGTTATTGGAGGAAGAGAGAGAAAAACGTTAAACAAATCAGACAGAAGCTGAGCTAGCAACCCGAGA
AGGCATCTATGAGAGGCCGTGGACTACCTGCCTGAGAGGGATGTTACGAGAGCCTCTGTCGTGGGGAGGGTGTCAAACCTGAC
ACCCCGTAGACAGAAGAGGCTTTCTGTAGGTACCACCATGGCAACAGGGCCCCACAGCTGCTCATTGCCCCCTTCAAAGAGGA
GGACGAGTGGGACAGCCCGCACATCGTCAGGTACTACGATGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAA
ACCTAAACTTGCACGAGCCACCGTTCTGTGATCCCAAGACAGGAGTCTCACTGTGCGCAGCTACCGGGTTTCCAAAAGCTCCTG
25 GCTAGAGGAAGATGATGACCTGTGTGGCCCGAGTAAATCGTCGGATGCGAGCATATCAGGGTTAACAGTAAAGACTGCAGA
ATTGTTACAGGTTGCAAAATATGAGAGTGGGAGGACAGTATGAACCGCACTTCGACTTCTCTAGGAATGATGAGCGAGATACTTT
CAAGCATTTAGGGACGGGAATCGTGTGGCTACTTTCTTAACTACATGAGTGTGTAAGCTGGTGGTGCCACCGCTTCTCC
TGATCTGGGGCTGCAATTTGGCCTAAGAAGGGTACAGCTGTGTTCTGGTACAACCTCTTGGCGAGCGGGGAAGGTGACTACCG
AACAAGACATGCTGCTGCTGCTGTGCTGTGGCTGCAAGTGGGTCTCCAATAAGTGGTTCATGAAGAGGACAGGAGTCTCTT
30 GAGACCTTGTGGATCAACAGAAGTTGACTGACATCCTTTCTGTCTCTCCCTTCTGCTGCTTCCAGCCATGTGCAACGTGACA
GACACCTTTGTATGTTCTTGTATGTTCTATCAGGCTGATTTTGGAGAAATGAATGTTGTCTGGAGCAGAGGGAGACCATA
CTAGGGCGACTCCTGTGTGACTGAAGTCCAGCCCTTCCATTACGCTGTGCCATCCCTGGCCCCAAGGCTAGGATCAAAGTGG
CTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGTGCCTTGTACCTCAGGTGTTTATAGGTGTGAGATGTTTCACTGAACCAA
AGTTCTGATACCTTGTGTTTACATGTTGTTTATGGCATTTCTATCTATGTGGCTTTACCAAAAAATAAATGTCCCTACCAG
35 AAGCCTTAAA

351

MPAHLQDDISSSYTTTTTITAPPPGVLQNGGDKLETMPLYLEDDIRPDIKDDIYDPTYKDKEGSPKVEYVWRNIILMSLLHL
GALYGITLIPTCKFYTWLWGVFFYFVSALGITAGAHRLWSHRSYKARLPLRLFLIIANTMAFQNDVYEWARHRAHHPSETHA
DPHNSRRGFFFSHVWLLVRKHPAVKEKGSTLDLSDLEAEKLVMFQRRYKPGLLMMCFLPTLVWPWFWGETFQNSVFVATFL
40 RYAVVLNATWLVNSAAHLFGYRKYDKNISPRENIVSLGAVGEGFHNHHSFPYDYSASEYRWIINFNTFFIDWMAALGLTYDR
KKVSKAAILARIKRTGDGNYKSG

352

45 GACGGTACCCGTTGCCAGCTCTAGCCTTTAAATTCCTGGCTCGGGGACCTCCACGCACCGCGGTAGCGCCGACAACCAGCTA
CGGTGCAAGCGCGCGGGCTCAGCGCGTACCGGGGGTTTCGAAACCGCAGTCCCTCCGGCGACCCGAACTCCGCTCCGGAGCC
TCAGCCCCCTGAAAGTGATCCCGGCATCGGAGAGCAAGATGCCGGCCCACTTGCTGCAGGACGATATCTCTAGCTCTATAC
CACCACCACCACATTACAGCGCTCCTCCAGGGGTCTGCAGAAATGGAGGAGATAAGTTGGAGACGATGCCCCCTACTTGGGA
AGACGACATTCGCCCTGATATAAAGATGATATATGACCCCACTACAAGGATAAGGAAGGCCCAAGCCCCAAGGTTGAATA
TGCTGAGAGAAACATCATCCTTATGTCTGTCTACACTTGGGAGCCCTGTATGGGATCACTTTGATTCTACTGCAAGTTCTA

5 GTTGCGCATCCAGGTGAGTCGCGTCAGCATGCACGACTATGAGGCCCTGCACTACGACAAGGAGCAGCTCAAGGAGGCCTCTGT
GCCGCTGGGCACCTGTGGTGGTCCCAGGAGACAGTGACCTAGAGCTCTGCCGTGCCACATTGAGAGACTCCAGCAGGAGCCCGA
TGGTGTGAGGCAAGTCCCCGACATGCCAGAACTGTCCCCAAGTGGTGTCTCTGGACGCCACCACTGCCAGCCGCTTCTA
CAGGATCGACCGAGCCAGGAGCACCTCAACTATGTGACTGAGATCGCACAGGATGAGATTTATATCTTGGACCTGAGCTGT
10 GGGGGCATCGGCCCGGCCCTGACCTCCCAACCCCCACTTCCCCCTCCCCACCTCACCTGTCTACCCACGCCCGGTCACCTGCA
AGGGGATGCTGCACCCCTCAAGGTGAAGAGCTGATTGAGGTGCCAAGAGGAACGACTTCTGTAACTCCAGGAGCTGCACCG
AGCTGGGGCGACCTCATGCACCGAGACGAGCAGAGTGCACGCTCTGCACACGCACTGAGCACTGGCAGCAAGGATGTGGT
CCGTACCTGTGGACACGCCCCCAGAGATCTTGTATGCGGTGGAGGAAACGGGGAGACCTGTTTGACCAAGCAGCGGC
CCTGGGCCAGCGCACCATCTGCCACTACATCGTGGAGGCCGGGGCTCGCTCATGAAGACAGACCAGCAGGGCGACACTCCCCG
15 GCAGCGGCTGAGAAGGCTCAGGACACGAGCTGGCCGCTTACCTGGAGAACCAGGAGCACTACCAGATGATCCAGCGGGAGGA
CCAGGAGACGGCTGTGTAGCGGGCCGCCACGGGCAGCAGGAGGACAATGCGGCCAGGGGACGAGCGCTTCTTGGCCACCT
CACTGCCACATTTCCAGTGGGACGGCCACGGGGGACCTAGGCCCCAGGGAAGAGCCCCATGCCGCCCTTAAAGAGCCGCCCA
GACTTAGGCTGGACTCAGGAGCTGGGGGGGCTTACCTGTTCCTTGGAGACCCCGCGGACCCGAGGCTCAGGGGAACAA
GACACGGCTGGGTGGATATGCCCTTGGCGGGTCTGGGGCAGGGCGCTCCCTGGCCGAGCAGATGCCCTCCAGGAGTGGGA
20 GGGGCTGGAGAGGGGAGGCTTCGGGAAGAGGCTTCTGGGCCCTGGTCTTGGCCGGGTCCCCAGCCCCGCTCTGCCCC
CACCCACCTCTCCGGGCTTCTCCCGAACTCAGCGCTGTGCACCTGCTTGGCTGCTTGGCTTGGCTTGGCTTGGCTTGGCT
15 GACCTCCCCGCTCCCCGTGATTTTCATCGCGGACTGTGCGGCTGGGGTGGGGGCGGGACTCTCACGGTGACATGTTTACA
GCTGGGTGTGACTCAGTAAAGTGGATTTTTTTTCTTTAAAAAAA
355

25 MQIPRAALLPLLLLLLAAPASQLSRAGRSAPLAAGCPDRCEPARCPPQPEHCEGGRARDACGCCCEVCGAPEGAACGLQEGPCG
EGLQCVVPFVGPASATVRRRAQAGLCVCASSEPVCGSDANTYANLQCLRAASRRSERLHRPPVIVLQRCAGCQGGEDPNSLRHK
YNFIADVVEKIAPAVVHIELFRKLPFSKREVPVAVSGSFVSEDGLIVTNAHVVTNKHVRKVELKNGATYEAKIKDVDEKADIA
LIKIDHQKLPVLLLRSSSELRPGEFVVAIGSPFSLQNTVTTGIVSTTQRGGKELGLRNSDMYIQTDAIINYNSGGPLVNLNLD
GEVIGINTLKVTAGISFAIPSDKIKKFLTESHDRQAKGAKITKKYIGIRMSLTSSKAKELKDRHRDPDVISGAYIIEVIPD
25 TPAEAGGLKENDVIISINGQSVVSANDVSDVIKRESTLNMVVRGNEDIMITVIPEEIDP
356

30 CCGGCCCTCGCCCTGTCCGCGCCACCGCCGCGCCGAGAGTGCCTATGCAGATCCCGCGCGCGCTCTTCTCCCGCTGCTG
CTGCTGCTGCTGGCGCGCCCGCTCGGCGCAGCTGTCCCGGGCGGCCGCTCGGCGCTTGGCGCGCGGGTGCACAGCCGC
TGCGAGCCGCGCGCTGCCCGCGCAGCCGAGCACTGCAGAGGCGCGCGGCCCGGGACGCGTGGCGCTGCTGCGAGGTGTGC
35 GCGCGCGCCGAGGGCGCGCGTGCAGGAGGCGCGCGTGCAGGAGGCGCGCGTGCAGTGCCTGCTGCGAGGTGTGC
GCCTCGGCCACGGTGGCGCGCGCGCAGGCGCGCTCTGTGTGCGCGCAGCAGCGAGCGGTGTGCGCGCAGCAGCCAAC
ACCTACGCCAACCTGTGCCAGCTGCGCGCGCCAGCGCGCTCCGAGAGGCTGCACCGCGCGCGGTCTGCTCTGCAGCGC
GGAGCTGCGGCCAAGGGCAGGAAGATCCCAACAGTTTGCGCCATAAATAAATTTATCGCGGACGTGGTGGAGAAGATCGCC
CCTGCGGTGGTTATATCGAATTGTTTCGCAAGCTTCCGTTTCTTAAACGAGAGGTGCCGCTGGCTAGTGGGTCTGGGTATT
40 GTGTCGGAAGATGGACTGATCGTGACAAATGCCCCAGTGGTGACCAACAAGCACCGGTCAAAGTTGAGCTGAAGAACGGTGCC
ACTTACGAAGCCAAAATCAAGGATGTGGATGAGAAAGCAGACATCGCACTCATCAAAATTGACCACCAGGCAAGCTGCCTGTC
CTGCTGCTTGGCCGCTCTCAGAGCTGCGGCGGGAGAGTTCGTGGTGGCATCGGAAGCCGTTTTCCTTCAAAACACAGTC
ACCACCGGGATCGTGAGCACCACCGAGCGAGGCGGCAAGAGCTGGGGCTCCGCAACTCAGACATGGACTACATCCAGACCGAC
GCCATCATCAACTATGGAACTCGGGAGGCGCGTTAGTAAACCTGGACGGTGAAGTGATTTGAAATTAACACTTTGAAAGTGACA
45 GCTGGAATCTCTTTGCAATCCCATCTGATAAGATTAAAGATTCTCAGGAGTCCCATGACCAGCAGGCCAAAGGAAAGCC
ATCACAAGAAGAAGTATATTGGTATCCGAATGATGTCACGTCCAGCAAAGCAAAGAGCTGAAGGACCGGCACCGGGAC
TTCCAGACGTGATCTCAGGAGCGTATATAATTGAAGTAATTCCTGATACCCAGCAGAAAGCTGGTGGTCTCAAGGAAACGAC
GTCATAATCAGCATCAATGGACAGTCCGTGGTCTCCGCCAATGATGTCAGCGACGTCATTAAAGGAAAGCACCTGAAATG
GTGGTCCGCGAGGGTAAATGAAGATATCATGATCAGAGTGAATTCGGAAGAAATTGACCCATAGGCAGAGGCATGAGCTGGACTT
45 CATGTTTCCCTCAAAGACTCTCCCGTGGATGACGGATGAGGACTCTGGGGTGGTGGAAATAGGACACTCAAGACTTTTACTGCT
ATTTTGTGTTTGTTCAGTGGAGACTCCCTGGCCAAAGAAATCTTCTGTAGATTGTCAGGCAAAACAAATGTAATGTTGAGATC
CGCAGGCGAAGCTCTGCCCTTCTGTATCTATGTATGCACTGTGCTTTTCTTGGCAGCTTGGGCCATCTTGTCTAGACAGT
CAGCATTTGTCTCTCTTTAACTGAGTATCATCTTGTGCAACTAATGCACTGATACAATGCGTAGATAGAAGAAGCCCCA
CGGGAGCCAGGATGGGACTGGTGTGTTTGTGCTTTTCTCAAGTCAGACCCAAAGGTCAATGCACAGAGACCCCGGTGGGT

GAGCGCTGGCTTCTCAAACGGCCGAAGTTGCCTCTTTTAGGAATCTCTTTGGAATTGGGAGCACGATGACTCTGAGTTTGAGCT
ATTAAAGTACTTCTTACAAA

357

MALETICRPSGRKSSKMQAFRIWDVNQKTFYLRNNQLVAGYLQGPVNVNLEEKIDVVPIEPHALFLGIHGGKMCVSKSGDETR
5 LQLEAVNITDLSENKQDKRFAFIRSDSGPTTSFESAACPGWFLCTAMEADQPVSILTNPDEGVMVTKFYFQEDE

358

AGCTCCACCCTGGGAGGGACTGTGGCCAGGTACTGCCCGGTGCTACTTTATGGGCAGCAGCTCAGTTGAGTTAGAGTCTGGA
AGACCTCAGAAGACCTCCTGTCTATGAGGCCCTCCCCATGGCTTTAGAGACGATCTGCCGACCCTCTGGGAGAAAATCCAGCA
AGATCAAGCCTTTCAGAATCTGGGATGTTAACCAGAAGACCTTCTATCTGAGGAACAACCAACTAGTTGCTGGATACTTGCAAG
10 GACCAAAATGTCAATTTAGAAGAAAAGATAGATGTGTACCATTTGAGCCTCATGCTCTGTTCTTTGGGAATCCATGGAGGGAAGA
TGTGCTGTCTGTGTCAAGTCTGGTGATGAGACCAGACTCCAGCTGGAGGCAGTTAACATCACTGACCTGAGCGAGAACAGAA
AGCAGGACAAGCGCTTCGCCTTCAATCCGCTCAGACAGTGGCCCCACCACAGTTTGTAGTCTGCCGCTGCCCGGTTGGTTCC
TCTGCACAGCGATGGAAGCTGACCAGCCCGTCAGCCTCACCAATATGCCTGACGAAGCGCTCATGGTCACCAAAATTTCTACTTCC
AGGAGGACGAGTAG

15 359

MIPNGYLMFEDENFIESSVAKLNALRKSGQPCDVRQVCGHEMLAHRVLAACSPYLFEIFNSDSPHGISHVKFDDLNPFAVE
VLLNYAYTAQLKADKELVKDVYSAKELKMDRVKQVCGDYLLSRMDVTSCISYRNFASCMGDSHLLNKVDAYIQEHLQISEEE
EFLKLPRLKLEVMLEDNVCLPSNGKLYTKVINWVQRSIWENGDSLEELMEEVQTLYYSDHKLLDGNLLDQAEVFGSDDDHQ
FVQKKPPRENGHKQISSSSSTGCLSSPNATVQSPKHEWKIVASEKTSNNTYLCLAVLDGIFCVIFLHGRNSPQSSPTSTPKLSKS
20 LSFEMQODELIEKPMSPMQUARSGLTAEMNGKLLAAGGYNREELRTVECYNPHTDHWSFLAPMRTPRARFQMAVLMGQLYVV
GGSNGHSDDLSCGEMYDSNIDDWIPVELRTRNRCNAGVCALNGKLYIVGGSDPYGQKGLKNCDFDPVTKLWTSAPLNIRRHQ
SAVCELGGYLYIIGGAESWNCNTVERYNPENNTWTLIAPMNVARRGAGVAVLNGKLFVCGGFDGSHAISCVEMYDPTTRNEWKM
MGHMTSPRSNAGIATVGNITIYAVEDSMAMNF

360

AGTGTCTCCCGGTGCGCGGTGGAGGTCGGTCGCTCAGAGCTGCTGGGCGCAGTTTCTCCGCTGCTGCTTCGGCGCGGCTGTAT
CGGCGAGCGAGCGAGTTCCCGCGAGTTCTCGGTGGCGCTCCCCCTTCCTTTTCAGTCTCCACGACTGGCCCCCTCGTCTTCTAC
TTGACCGCTCCCGTCTTCTGCGCGCTTTCGGCGCTTTCGGTGGGCGGATTCCCGCCCGCTTCTCCTGCTTCCCATCGAAGCT
CTAGAAATGAATGTTTCCATCTCTTTCAGAGATGAACCAGATTATGATGCATCATATACAGAGAAGAAATTCGTGTCTATAGCTT
TTAAGGACTTGATTACATCATTTTCAAGCCTGATAGTTTGAATCACCATTAGAGCTTAAGACACACCTGCCCTTCATTTCAAC
30 CACCTGTCTTCATACCCTGACGAAGTGCACCTTTTAACACTCCTTTGTCTTGGATTACTTAAGAGTTCACAGAAATACATTTG
CCACCAACAGAGTAGCCTAAATTTATAAGGAAAATGATTCCCAATGGATATTTGATGTTTGGAGATGAAAATTTTATTGAGTCT
TCTGTTGCCAAATTTAATGCCCTGAGGAAAAGTGGCCAGTTCTGTGATGTTTCGACTTCAGGTCTGTGGCCATGAAATGTTAGCA
CACAGAGCAGTGCTAGCTTGTGTCAGTCCCTATTTATTTGAAATCTTTAATAGTGATAGTGATCCTCATGAAATTTCTCACGTT
AAATTTGATGATCTCAATCCAGAAGCTGTTGAAGTCTTGTGAATTTATGCCTACACTGCTCAGTTGAAAGCAGATAAGGAATTG
35 GTAAAAGATGTTTATTTCTGCAGCAAAAGAGCTGAAGATGGATCGAGTAAAGCAGGTTTGTGGTGATTTTACTGTCTAGAATG
GATGTTACAGCTGCATCTCTTACCGAAATTTGCAAGTTGATGGGAGACTCCCATTTGTTGAATAAGGTTGATGCTTATATT
CAGGAGCATTTGTTACAAATTTCTGAAGAGGAGGAGTTTCTTAAGCTTCCAAGGCTAAAGTTGGAGGTAATGCTTGAAGATAAT
GTTTGTCTGCCAGCAATGGCAAATTTATATACAAAGGTAATCAACTGGGTGCAGCGTAGCATCTGGGAGAAATGGAGACAGTCTG
GAAGAGCTGATGGAAGAGGTTCAAACCTTGTAATCTAGCTGATCACAAGCTGCTTGATGGGAACCTACTAGATGGACAGGCT
40 GAGGTGTTTGGCAGTGATGATGACCACATTAGTTTGTGCAGAAAAGCCACCACGTGAGAAATGGCCATAAGCAGATAAGTAGC
AGTTCAACTGGATGCTCTCTCTTCTCAAATGCTACAGTACAAAGCCCTAAGCATGAGTGGAATAATCGTTGCTTTCAGAAAAGACT
TCAAATAACACTTACTTGTGCTGGCTGTGCTGGATGGTATATTCTGTGTCATTTTCTTTCATGGGAGAAACAGCCACAGAGC
TCACCAACAAGTACTCCAAAATAAGTAAGAGTTTAAGCTTTGAGATGCAACAAGATGAGCTAATCGAAAAGCCCATGTCTCCT
ATGCAGTACGCACGATCTGGTCTGGGAACAGCAGAGATGAATGGCAAACCTCATAGCTGCAGGTGGCTATAACAGAGAGGAATGT
45 CTTGCAACAGTCGAATGCTATAATCCACATACAGATCACTGGTCTTTCTTGTCTCCCATGAGAACCAAGAGCCGATTTCAA
ATGGCTGTACTCATGGCCAGCTCTATGTGGTAGGTGGATCAAATGGCCACTCAGATGACCTGAGTTGTGGAGAGATGTATGAT
TCAAACATAGATGACTGGATTCTGTTCAGAAATGAGAACTAACCGTTGTAATGCAGGAGTGTGTGCTCTGAATGAAAAGTTA
TACATCGTTGGTGGCTCTGATCCATATGGTCAAAAAGGACTGAAAAATTTGTGATGTATTTGATCCTGTAACAAAGTTGTGGACA

AGCTGTGCCCTCTTAACATTCGGAGACACCAGTCTGCAGTCTGTGAGCTTGGTGGTTATTTGTACATAATCGGAGGTGCAGAA
TCTTGGAATTGTCTGAACACAGTAGAACGATACAATCCTGAAAATAATACCTGGACTTTAATTCACCCATGAATGTGGCTAGG
CGAGGAGCTGGAGTGGCTGTTCTTAATGGAAGAACTGTTGTATGTGGTGGCTTTGATGGTTCTCATGCCATCAGTTGTGTGGAA
ATGTAATGATCCAAC TAGAAATGAATGGAAGATGATGGGACATATGACTTCACCAAGGAGCAATGCTGGGATTGCAACTGTAGGG
5 AACACCATTATGCAGTGGAGGATTTCGATGGCAATGAATTTCTGAATACGGCGGAAGTCTATAACCTTGAGTCAAATGAATGGA
GCCCCATACAAAGATTTTCCAGTTTAAACGGGTTTAAAGACCTCTCAAAC TAACAGGCTTAGTGATGTAATATGTTAGCAG
AGGTACACTTGTGAATAAAGAGGGTGGTGGGGATAGATGTTGCTAACAGCAACACAAAGCTTTTGCATATTGCATACTATTAA
ACATGCTGTACATACTTTTGGGTTTATTTGGGAAGGAATGCAAAGATGAAGGTCTGTTTGTGTACTTTTAAGACTTTGGTTAT
TTTACTTTTGGAAAAGATAAACCAAGATTGATTGGGCACATCAAAAAAAAAAAAAAAAAAAAAA

10 361
MVDGVMILPVLIMIALPSPSMEDEKPKVNPPLYMCVCEGLSCGNEDHCEGQQCFSSLSINDGFHVYQKGCQVYEQGKMTCKTP
PSPGQAVECCQGDWCNRNITAQLPTKGSFPQTQNFHLEVGLIILSVVFAVCLLACLGLVALRKFRRNQERLNPRDVEYGTIE
GLITTNVGDSTLADLLDHSC TSGSGSLPFLVQRTVARQITLLECVKGGRYGEVWRGWSQGENVAVKIFSSRDEKSWFRETELY
NTVMLRHENILGFIASDMTSRHSSTQLWLITHYHEMGSYDYLQLTLDTVSCLRIVLSIASGLAHLHIEIFGTQGKPAIAHRD
15 LKSKNILVKNGQCCIADLGLAVMHSQSTNQLDVGNPRVGT KRYMAPEVLDETIQVDCFD SYKRVDIWAFLVLWEVARRMVS
NGIVEDYKPPFYDVVPNDPSFEDMRKVVCVDQQRPNIPNRWFS DPTLTSLAKLMKECWYQNP SARLTALRIKKT LT KIDNSLDK
LKTDC

362
GAAGCGAATAGCGTTTTTCAGAGATATTGGGCGGCTCAAGGGTCTTACTCTGTGCGCCAGTCTGTAATGCAGTGTGTGACCATA
20 GCCCAGTGCAGCTCCACCTCCAGGCTCAAGCAGTCTTCCCCCTCGCCCTCATGAATAGCTGGGACTACAGCTGGAGCAT
TGTAAGCGTCACTAGTCCAAAGTGAGAGCTGCTGGAGAACTCATATCCCAGGAACGCCTCTTCTACTCTCCGAGTACCCAG
TGACCAGAGTGAGAGAAGCTCTGAACGAGGGCAGCGGCTTGAAGGACTGTGGGCAGATGTGACCAAGAGCCTGCATTAAGTTG
TACAATGGTAGATGGAGTGATGATTCTTCCTGTGCTTATCATGATTGCTCTCCCTCCCTAGTATGGAAGATGAGAAGCCCAA
GGTCAACCCCAAACCTCTACATGTGTGTGTGAAGGTCTCTCCTGCGGTAATGAGGACCACTGTGAAGGCCAGCAGTGTCTTTTC
25 CTCAGTGAAGCATCAACGATGGCTTCCACGTCTACCAGAAAGGCTGCTTCCAGGTTTATGAGCAGGGAAGATGACCTGTAAGAC
CCCGCGTCCCTGGCCAAGCTGTGGAGTGTGCAAGGGGACTGGTGTAACAGGAACATCACGCCCCAGCTGCCCACTAAAGG
AAAATCCTTCCCTGGAACACAGAATTTCCACTTGGAGGTTGGCTCATTTATCTCTCTGTAGTGTTCGAGTATGCTTTTAGC
CTGCCTGTGGGAGTTGCTCTCCGAAAATTTAAAAGGGCGCAACCAAGAACGCCTCAATCCCGAGACGTGGAGTATGGCACTAT
CGAAGGGCTCATCACCAATGTGGAGACAGCACTTTAGCAGATTTATTGGATCATTCGTGTACATCAGGAAGTGCTCTGG
30 TCTTCTCTTTCTGGTACAAAGAACAGTGGCTCGCCAGATTACACTGTGGAGTGTGTGCGGAAAGGCAGGTATGGTGAGGTGTG
GAGGGGAGCTGGCAAGGGGAAAATGTTGCCGTGAAGATCTTCTCTCCCGTGATGAGAAGTCATGGTTACAGGAAACGGAATT
GTACAACACTGTGATGCTGAGGCATGAAAATATCTTAGGTTTCATTGCTTCAGACATGACATCAAGACACTCCAGTACCCAGCT
GTGGTTAATTACACATTATCATGAAATGGGATCGTTGTACGACTATCTTCAGCTTACTACTCTGGATACAGTTAGCTGCCTTCG
AATAGTGTGTCCATAGCTAGTGGTCTTGCACATTTGCACATAGAGATATTTGGGACCAAGGGAAACAGCCATTGCCCATCG
35 AGATTTAAAGAGCAAAAATATTCTGGTTAAGAAGAATGGACAGTGTTCATAGCAGATTTGGGCTTGGCAGTATGCATTCCCA
GAGCACCATCAGCTTGATGTGGGGAACAATCCCGTGTGGGCACCAAGCGCTACATGGCCCCGGAAGTTCTAGATGAAACCAT
CCAGGTGGATTGTTTCGATTCTTATAAAAGGGTCGATATTTGGGCTTTGGACTTGTTTTGTGGGAAGTGGCCAGGCGGATGGT
GAGCAATGGTATAGTGGAGGATTACAAGCCACCGTTCTACGATGTGGTTCCCAATGACCCAAGTTTGAAGATATGAGGAAGGT
AGTCTGTGTGGATCAACAAAGGCCAAACATACCCAACAGATGGTTCTCAGACCCGACATTAACCTCTCTGGCCAAGCTAATGAA
40 AGAATGCTGGTATCAAAATCCATCCGCAAGACTCACAGCACTGCGTATCAAAAAGACTTTGACCAAAAATTGATAATTCCTCGA
CAAAATTGAAAAC TGACTGTTGACATTTTCATAGTGTCAAGAAGGAAGATTTGACGTTGTTGTCATTGTCAGCTGGGACCTAAT
GCTGGCTGACTGGTTGTGCAATGGAATCCATCTGTCTCCCTCCCAAAATGGCTGCTTTGACAAGGCAGACGTCGTACCCAGC
CATGTGTGGGGAGACATCAAAACCACCTAACCTCGCTCGATGACTGTGAAC TGGGCATTTACCGAACTGTTACACTGCAGA
GACTAATGTTGGACAGACACTGTTGCAAAAGGTAGGACTGGAGGAACACAGAGAAATCCTAAAGAGATCTGGGCATTAAGTCA
45 GTGGCTTTGCATAGCTTTTCAAGTCTCTTAGACTCTCCACGGGAAACTCAAGGAGGTGGTGAATTTTAAATCAGCAATATT
GCTGTGTGCTCTCTTCTTTATTGCACTAGGAATTTCTTTGCATTCCTTTACTTGTGCACTGTTACTCTTAAATTTAAAGACCCAAT
GCCAAAATGTTGGCTGCGTACTCCACTGGTCTGTCTTTGGATAATAGGAATTC AATTTGGCAAAAACAAATGTAATGTGCACT
TTGTGCAATTTTACACATGTGCTGATGTTTACAATGATGCCGAACATTAGGAATTTGTTTATACACAATTTGCAAAATTTAT
TACTTGTGCACTTAGTAGTTTTTACAAAAC TGTGTCATATGTTAAAGCTTATTTTATGTGGTCTTATGATTTTATTACA
50 GAAATGTTTTAACACTATACTCTAAAATGGACATTTTCTTTTATATCAGTTAAAATCACATTTTAAGTGTTCACATTTGTA

TGTGTGTAGACTGTAACCTTTTTTTCAGTTCATATGCAGAACGTATTTAGCCATTACCCACGTGACACCACCGAATATATTATCG
ATTTAGAACCAAAGATTTCAGTAGAATTTAGTCCTGAACGCTACGGGGAAAAATGCATTTTCTTCAGAAATTACCATACGTGC
ATTTAAACTCTGCCAGAAAAAATAACTATTTTGTGTTTAACTACTTTTGTATTTAGTAGTTATTTGTATAAATTAATAAAC
TGTTTTCAAGTC

5 363

MYQVPLPLDRDGLVRLRFTMVALVTCCPLVAFILWSLLFHFKEETTATHCGATPCRMFSAASQPLDPDGLFRLRFTAMV
WWAITFPVFGFFFCIIWSLVFHFYETVATDCGVPNYLPSVSSAIGGEVPQRYVWRFCIGLHSAPRFLVAFAYWNHYLSCTSPCS
CYRPLCRLNFGNLNVVENLALLLVLYVSSSEDFTHENAFIVFIASSLGHMLLTCLWRLTKKHTVSQEDRKSYSWKQRLFIINF
ISFFSALAVYFRHNMCEAGVYTI FAILEYTVVLTNMAFHMTAWWDFGNKELLITSQPEEKR

10 364

GGCCTTACCAATCGCGAAAAACCCGCCGTTTCGCGCTCTGACCAGCCCGCAGAGCCAGCCCCGACCCCGGGCCACCTGGGCCCC
GGGTTCCGCCGGCACTCTCGCCACCACCGCGTGGGTCTGACAAGATGTACCAGGTCCCACTACCACTGGATCGGGATGGGACCC
TGGTACGGCTCCGCTTCACCATGGTGGCCCTGGTCACGGTCTGCTGTCCACTTGTGCGCTTCCTCTTCTGCATCCTCTGGTCCC
TGCTCTTCCACTTCAAGGAGACAACGGCCACACACTGTGGGGCCACGCCCTGCAGGATGTTCTCTGCGGCTCCAGCCTTTGG
15 ACCCCGATGGGACCTTGTTCGGCTTCGCTTCACAGCCATGGTCTGGTGGGCCATCACTTTTCTGTGTTCCGGCTTCTTCTTCT
GCATCATCTGGTCCCTGGTGTTCACCTTTGAGTACACGGTGGCCACTGACTGTGGGGTCCCAATTACCTGCCCTCGGTGAGCT
CAGCCATCGCGGGGAGGTGCCCCAGCGTACGTGTGGCGTTTCTGCATCGGCTGCCTCGCGCCTCGCTTCTTGGTGGCCT
TCGCTTACTGGAACCACTACCTCAGCTGCACCTCCCGTGTCTCTGCTATCGCCCGCTCTGCCGCTCACTTCGGCCTCAATG
TCGTGGAGAACCTCGCGTGTCTAGTGTCTCACTTATGTCTCCTCCTCCGAGGACTTCACCATCCACGAAATGCTTTCATTTGTGT
20 TCATTGCCCTCATCCCTCGGGCACATGCTCCTCACCTGCATTCTCTGGCGGTTGACCAAGAAGCACACAGTAAGTCAGGAGGATC
GCAAGTCTACAGCTGGAACAGCGGCTCTTCATCATCACTTCATCTCCTTCTCTCGGCGCTGGCTGTCTACTTTCCGCACA
ACATGTATTGTGAGGCTGGAGTGTACACCATCTTTGCCATCTCGGAGTACACTGTTGTCTTAACCAACATGGCGTTCCACATGA
CGGCCTGGTGGGACTTCGGGAACAAGGAGCTGCTCATAACCTCTCAGCCTGAGGAAAAGCGATTCTGAACCTTCAGTCTGTCT
TGGGAGGACGACGCCACTGCCCAGAAACAAGAACACGATACCATTTGCGCTTCCCCACCCACATCCTCTCTTGGCCTTAC
25 TGAAGATGGGGGAAGGGTAAGAAGGAAGGGTGTAGGCCAAGGCTCACCCAGTGTCTGGCTTCTCTCTCCACCCCTCATAT
GGCGTGGGGTCTCAACATCACCTTTACCTGAGAGGCCCAAGAAGCTGAGCTGGCAGAGAGCTCCACCATTTGGTGTCTAAA
AAAAAAACGTCCTGAGGTTTCATGACCACCATCCAGTTTCTGGCCTTTACACAGTCACTTTTCACTGAGGTGAGGAGCCCTGA
GCAGTGGCTGTCTCCCTGACAACACAGCCATTCTCTGCACGGGGTTCATTATAGGACTAATGTATTTTCATGATCTACTGTGC
ACATCCAGGCTGTGGCCACAGTCCCTGCTAAAGTTGCTCAGGTGTTCTAGTCTGACTTCACCTTTTTCATTGGTGTGTGC
30 CCTAGGGTATGTACCTTCCCCATCTGAGCCTCGGTGTGTCATGTGTCTGGCGGGGATGGGTGGACTGTATGATTTCCAAGG
ACTCTACCAGTCACTGGTCTGATGTCATCGGTGGAGGTGGTGTCTATACCTAAAGGATGACCTGTCTCCAGAAACAGACCA
GCACAGCATGTATTTCTCTCTTCTGAAAGTTCTGGCTTGTAGACCCCTCCCTCCTTTGCAAAGGTATGGGATAGAGGGGTC
AGATGCAGATCTCTACTGTAAATGGGCTCCCTGGTATCTCTGTCTTCCCTACTGTCTCAAACCTTAAATTTTGGTTGTACAT
TTTATTTTGAAGGAAATAAATTTTCTTTTGGGCCAACAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
35 AAAAAAAAAAAAAAAAAAAAAA

365

MMLSLNNLQNIYNPVIFVGTIPDQLDPGLIVIRGHVPSDADRQVDLQNGSSMKPRADVAFHFNPRFKRAGCIVCNTLINE
KWGREITYDTPFQKEKKSFEIVIMVLKAKFQVAVNGKHTLLYGRIGPEKIDTLGIYGVNIHSIGFSFSSDLQSTQASSLEL
TEISRENVKSGTPQLRLPFAARLNTPMGPGRTVVVKGEVNANAKSFNVDLLAGKSKDIALHLNPRNLKAFVRNSFLQESWGE
40 EERNITSPFPSPGMVFEMIIYCDVREFKAVNGVHSLEYKHFELSSIDTLEINGDIHLLEVRW

366

ACACAGAAGAGACTCCAATCGACAAGAAGCTGGAAGAAGATGATGTTGTCTTAAACAACCTACAGAATATCATCTATAACCCG
GTAATCCCGTTTGTGTCACCATCTCTGATCAGCTGGATCCTGGAACCTTTGATTGTGATACGTGGGCATGTTCTTAGTGACGCA
GACAGATTCAGGTGGATCTGCAGAAATGGCAGCAGCATGAAACCTCGAGCCGATGTGGCCTTTCAATTCATCTCGTTTCAAA
45 AGGGCCGGTGCATTTGCAATACTTTGATAAATGAAAAATGGGGACGGGAAGAGATCACCTATGACACGCCCTTCCAAAAA
GAGAAAAAGTCTTTGAGATCGTGATTATGGTGTGAAGGCCAAATTCAGGTGGCTGTAAATGGAAAAACATCTGTCTCTAT
GGCCACAGGATCGGCCAGAGAAAAAGACACTCTGGGCATTTATGGCAAAGTGAATATTAATCAATTTGGTTTGTAGCTTCAGC
TCGGACTTACAAAGTACCAAGCATCTAGTCTGGAAGTACAGAGATAAGTAGAGAAATGTTCCAAAGTCTGGCAGCCCCAG

10 MAALTRDPQFQKLQQWYREHRSELNLRRLFDANKDRFNHFSLTNTNHHGILVDYSKNLVTEDVMRMLVDLAKSRGVEAARERM
FNGEKINYTEGRAVLHVALNRNSENTPIILVDGKDVMPENVKVLDKMKSFCQQRVRSGDWKGYTGKTTITDVINIGIVGSDLGPLMVT
EALKPYSSGGPRVWVYSNIDGTHIAKTLAQLNPESSLFIASKTFTTQETITNAETAKEWFLQAAKDPSAVAKHFVALSTNTTK
VKEFGIDPQNMFEFWDWVGGRYLSWSAIGLSIALHVGFDNFEQLLSGAHWMDDQHFRITTPLEKNAPVLLALLGIWYINCFGCETH
AMLPIYDLNHLRFAAYFQQQDMESNGKYITKSGTRVDHTGPIVWGEPTNGQHPAFYQLIHQGTKMIPCDFLI PVQQTQHPIRKGL-
HHKLLQFLAQTAELEMRGKSTAEARKELQAAGKSPEDLERLLPHKVFEGNRPTNSIVPTKLTFFMLGALVAMYEHKIFVQGI I
15 WDINSFDQGWELGKQLAKKIEPELDGSAQVTSHDASTNGLINFIKQQRREARVQ

CTCAGAGACTCCGCCATGGCCGCTCTACCCGGGACCCCAAGTTCCAGAAGCTGCAGCAATGGTACCGCGAGCACCGCTCCGAG
CTGAACCTGCGCCGCCTCTTCGATGCCAACAAGGACCGCTTCAACCATTTCAGCTTGACCCTCAACACCAACCATGGGCATATC
CTGGTGGATTACTCCAAGAACCCTGGTGACGGAGGACGTGATGCGGATGCTGGTGGACTTGGCCAAGTCCAGGGGCGTGGAGGCC
GCCCGGGAGCGGATGTTCAATGGTGAGAAGATCAACTACACCGAGGGTCGAGCCGTGCTGCACGTGGCTCTGCGGAACCGGTCA
AACACACCCATCTTGGTAGACGGCAAGGATGTGATGCCAGAGGTCAACAAGTTCTTGGACAAGATGAAGTCTTTCTGCCAGCGT
GTCCGGAGCGGTGACTGGAAGGGGTACACAGGCAAGACCATCACGGACGTCATCAACAATTGGCATTGTGCGCTCCGACCTGGGA
CCCCTCATGGTGAAGCCCTTAAGCCATATCTTTCAGGAGTCCCGCGTCTGGTATGTCTCCAACATTGATGGAACCTCAC
ATTGCCAAACCTTGGCCAGCTGAACCCGAGTCCTCCTGTTTCATCATTTGCCCTCCAAGACCTTTACTACCCAGGAGACCATC
ACGAATGCAGAGACGGCGAAGGAGTGGTTTCTCCAGGCGGCCAAGGATCCTTCTGCAGTGGCGAAGCACTTTGTGTGCCCTGTCT
ACTAACACAACCAAAGTGAAGGAGTTTGGAAATTGACCCCTCAAAACATGTTTCGAGTTCTTGGGATTGGGTGGGAGGACGCTACTCG
CTGTGGTTCGGCCATCGGACTCTCCATTGCCCTGCACGTGGGTTTGTGACAACTTCGAGCAGCTGCTCTCGGGGGCTCACTGGATG
GACCAGCACTTCCGCACGACGCCCTTGGAGAAGAACGCCCCGCTTGTCTGGCCCTGCTGGGTATCTGGTACATCAACTGCTTT
GGGTGTGAGACACACGCCATGCTGCCCTATGACCAGTACCTGCACCGCTTGTCTGCGTACTTCCAGCAGGGCGACATGGAGTCC
AATGGGAAATACATCACCAAATCTGGAACCCGTGTGGACCACCAGACAGGCCCCATTGTGTGGGGGGAGCCAGGGACCAATGGC
CAGCATGCTTTTTACCAGCTCATCCACCAAGGCACCAAGATGATACCTGTGACTTCCCTCATCCCGGTCCAGACCCAGCACCCC
ATACGGAAGGGTCTGCATCACAAGATCTCTCTGGCCAACTTCTTGGCCAGACAGAGGCCCTGATGAGGGGAAAATCGACGGAG
GAGGCCGAAAGGAGCTCCAGGCTGCGGGCAAGAGTCCAGAGGACCTTGAGAGGCTGCTGCCACATAAGGTCTTTGAAGGAAAT
CGCCCAACCAACTCTATTGTGTTTACCAAGCTCACACCAATTCATGCTTGGAGCCTTGGTGCCTATGTATGAGCACAAGATCTTCT
GTTTCAGGGCATCATCTGGGACATCAACAGCTTTGACCAGTGGGGAGTGGAGCTGGGAAAGCAGCTGGCTAAGAAAAATAGAGCCT
GAGCTTGATGGCAGTGCTCAAGTGACCTTCCAGCAGCTTCTACCAATTGGGCTCATCAACTTCATCAAGCAGCGCGGAGCC
AGAGTCCAATAAACTCGTGCTCATCTGACGACCTCTCTGTGACTCCCCCTTTCTCTCTGTCGTCCTCTCTCCCGGAGCCGCGACT
GCATGTTCTTGGACACCACCGAGCACCCTCTGGTGTGGGCTTGGACACAGCCCTTTCAGAGGGAAGGCTGGTCTCCCCCA
GCCTAACCCCCAGCCCCCTCATGTCTATGCTCCCTCTGTGTTAGAATTGGCTGAAGTGTTTTGTGTCAGCTGACTTTTCTGACC
CATGTTACAGTTGTTTCACATCCCATGTAGAAAAACAAAGATGCCACGGAGGAGGT

45 MKKEHVLHCQFSAWYPPFRGVTIKSVILPLPQNVKDYLLDDGTLVVSGRDDPPTHSQPDSDEAEETIQWSDDENTATLTAPEFP
EFATKVQEPINSLGGSVFPKLNSAPRAYWIAMNSCLKCTLSIDIFLLPKSSDFITRDFTQPFICHDTDDSPDCIEYELVLRK
WCELIPGAEFRFCVKENKLIGISQRDYTQYYDHISKQKEEIRRCIQDFFKKHIQYKFLDEDFVFDIYDRSGKVLWIDFNPFGE
VTDSSLFTWEELISENNLNGDFSEVDAQEQDSPAFRCTNSEVTVPSPYLSYRLPKDFVDLSTGRDAHKLIDFLKLRNQOEDD

TCGCGTTTAGGGCGAAGACGGAGTTGTAACCTTCTTAAATTCCTCTCTCGACACTTCGGTAATTCCTCTTTTCGAGACTAAAGCT
CTTTTGTATGCGTGTGTGTCAAGCGTATGCCCGGGATTCTCCTCCGCTTCCTTTTCTCGGTCTTCCTCTTGCCTTAGGGAC

- CGGAAGAGTCCTTGAACCAAAATAGCTCGGCGGGCACTTCCGGGGCCGGCGCCAGAGTTCGGGAGGGTGCAGGCAGGAGAGG
GAAAGGCAGCAGCGCGGCGAGCTGGAGGATGAAGAAGGAGCATGTGCTTCACTGCCAGTTCTCCGCGTGGTACCCGTTCTTCCG
AGGCGTTACCATCAAGAGTGTCATTTCTTCCACTTCTCAGAATGTGAAGGATTATTTACTCGATGATGGAACCTCTGGTGGTTTC
AGGAAGGGATGATCCACCAACACATTTCTCAGCCAGACAGTGATGATGAAGCAGAAGAAATACAGTGGTCTGATGATGAGAACAC
5 AGCCACGCTTACGGCACCAGAATTTCTTGAGTTTGCCACTAAAGTCCAGGAACCTATCAATTCCTCGGGGGCAGTGTCTTTCC
TAAGCTTAATTGGAGTGCCCCAAGGGATGCGTATTGGATAGCAATGAATAGTTCTCTGAAATGTAAAACCCCTCAGCGACATCTT
CTGCTTTTCAAGAGTTCCGATTTTCATCACTCGTGACTTCACTCAGCCGTTTATTTCATTGTACTGATGATTTCCAGATCCATG
TATAGAATATGAGCTCGTTCTCCGAAAATGGTGTAATTGATTCTTGGGGCTGAGTTTCGATGTTTTGTCAAGGAAAACAAGCT
TATTGGTATTTCTCAAAGAGACTACACACAATACTATGATCATATTTCTAAACAAAAGGAAGAAATTCGCAGATGCATACAAGA
10 CTTTTTCAAGAAACACATACAGTACAAATTTCTAGATGAAGACTTTGTGTTTCGATATATACAGAGACAGTAGGGGAAGGTGTG
GCTCATTGACTTTAATCCATTTGGTGAAGTCACAGATTCCTGCTGTTTACCTGGGAAGAACTGATATCTGAGAAACAACTTAA
CGGCGATTTTAGTGAAGTTGACGCTCAAGAGCAGGATTTCCAGCTTTCCGTTGCACAAACAGTGAAGTGACAGTCCAGCCCAG
CCCCATTGTGAGTTACCGGCTACCAAGGACTTTGTAGACCTCTCTACTGGGAGGGACGCTCACAAGCTAATAGACTTCCCTTAA
GCTGAAGAGAAATCAGCAGGAGGACGACTGATGAGCGTACTGTAACTGGAGAAGAGGAGGCCCCCACCCTCCGGGAGCT
15 GCTCATCAGCCGCAACTTCTGCGGACCTGATGCGGGTGGGCCAGCAGTGTGGACATCAGCCACTTTTATATTATGATGTACA
TTCACCTGGGGAACCAACCGAGGAGCTTTGCTACTTGTAAAAATAACATAATAAATAGATCTTAAACATAGGAAAACCATACT
GTTCTGATAATAAAATGCTTTCTATGAAAT
- 371
- MERIPSAQPPACLPKAPGLEHGDLPMPYPAHMYQVYKSRRGIKRSEDSKETYLPHRLIEKKRRDRINECIAQLKDLLPEHLK
20 LTTLGHLEKAVVLELTLKHVKALTNLIDQQQKIIALQSLQAGELSGRNVETGQEMFCSGFQTCAREVLQYLAKHENTRDLKS
SQLVTHLHRVVELLQGGTSRKPSDPAPKVMDFKEKPSSPAKGSBGPKNCPVVIQRTFAHSSGBQSGSDTDTDSGYGGESEKG
DLRSEQPCFKSDHGRRTMGERIGAIIKQSEEPPTKKNRQLSDDEGHFTSSDLISSPFLGPHPHQPPCLPFYLIPPSATAYL
PMLEKCWYPTSPVLYPGLNASAAALSSPMNPKISAPLMPQRLPSPLPAHPSVDSSVLLQALKPIPLNLETKD
- 372
- GGACACCGGGCCATGCACGCCCCAACTGAAGCTGCATCTCAAAGCCGAAGATTCCAGCAGCCCAGGGGATTTCAAAGAGCTCA
GACTCAGAGGAACATCTGCGGAGAGACCCCCGAAGCCCTCTCCAGGGCAGTCTCATCCAGACGCTCCGCTAGTGACAGACAGGA
GCGCGCAGTGGCCCGGCTCGCCGCGCCATGGAGCGGATCCCCAGCGCGCAACCACCCCCCGCTGCCGCCAAAGCAGCCGG
ACTGGAGCACGGAGACCTACCAGGGATGTACCCCTGCCACATGTACCAAGTGTAAGTCAAGACGGGAATAAAGCGGAGCGGA
GGACAGCAAGGAGACCTACAAATTGCGCGACCGGCTCATCGAGAAAAAGAGACGTGACCGGATTAAACGAGTGCATCGCCAGCT
30 GAAGGATCTCTACCCGAACATCTCAAACCTTACAACCTTTGGGTCACTTGGAAAAAGCAGTGGTTCTTGAACCTTACCTTGAAGCA
TGTGAAAGCACTAACAAACCTAATTGATCAGCAGCAGCAGAAAAATCATTGCCCTGCAGAGTGGTTTACAAGCTGGTGAGCTGTC
AGGGAGAAATGTGAAACAGGTCAAGAGATGTTCTGCTCAGGTTTCCAGACATGTGCCCGGAGGTGCTTCAGTATCTGGCCAA
GCACGAGAACACTCGGACCTGAAGTCTTTCGAGCTTGTACCCACCTCCACCGGGTGGTCTCGGAGCTGCTGCAGGGTGGTAC
CTCCAGGAAGCCATCAGACCCAGCTCCCAAAGTGATGGACTTCAAGGAAAAACCCAGCTCTCCGGCCAAAGGTTCCGAAGGTCC
35 TGGGAAAAACTGCGTGCCAGTCTCCAGCGGACTTTCGCTCACTCGAGTGGGAGCAGAGCGGCAGCGACACGACACAGACAG
TGGCTATGGAGGAGAAATCGGAGAGGGCGACTTTCGCGAGTGAGCAGCCGTGCTTCAAAAGTGACCACGAGCGCAGGTTACGAT
GGGAGAAAGGATCGCGCAATTAAGCAAGAGTCCGAAGAACCCCCACAAAAAGAACCGGATGCAGCTTTCGGATGATGAAGG
CCATTTCACTAGCAGTGACCTGATCAGCTCCCCGTTCTTGGGCCACACCCACACCCAGCTCTTCTGCTGCGCTTCTACCT
GATCCACCTTCAGCGACTGCCTACCTGCCCATGCTGGAGAAGTGCTGGTATCCACCTCAGTGCCAGTGCTATACCCAGGCCCT
40 CAACGCTCTGCGCGAGCCCTCTCTAGCTTCATGAACCCAGACAAGATCTCGGCTCCCTTGGCTCATGCCCCAGAGACTCCCTTC
TCCCTTGCAGCTCATCCGTCCGTGACTTCTGTCTTGTCTCCAAGCTCTGAAGCCAATCCCCCTTTAAACTTAGAAACCAA
AGACTAACTCTCTAGGGGATCTGCTGCTTTGCTTCTCTCGCTACTTCTTAAAAAGCAACAAAAAGTTTGTGTAATG
CTGCAAGATTGTTGATTGTGTATACTGAGATAATCTGAGGCATGGAGAGCAGATTCAAGGTGTGTGTGTGTGTGTGTGTGT
GTGTGTATGTGCGTGTGCGTGACATGTGTGCTGCTGCTGTTGGTATAGGACTTTAAAGCTCCTTTTGGCATAGGGAAGTCACGA
45 AGGATTGCTTGGACATCAGGAGACTTGGGGGGGATTGTAGCAGAGCTGCGGCTTTTCCACCCAGAGATAGCCCCCTTCGAT
ACACATCAGCTGGATTTTCAAAGCTTCAAAGTCTTGGTCTGTGAGTCACTTTCAGTTTGGGAGCTGGGTCTGTGGCTTTGAT
CAGAAGGTACTTTCAAAGAGGGCTTTCCAGGGCTCAGCTCCCAACAGCTGTTAGGACCCACCTTTTGCCTTTATTGTGCGA
CGTGACTCACCAGAGCTCGGGGAGAGAGCAGTCAAGCCAGCTTTCTGCTAACATGGGGAGGTAGCAGGCACTGGCATAGCA
CGGTAGTGGTTTGGGGAGGTTCGCGAGGTCTGCTCCCCACCCCTGCTCGGAAGAATAAAGAGAATGTAGTTCCTTACTCAGG
50 CTTTCTAGTGATTAGCTTACTAAGGAAGTGAATAAGGGCCCTTGTACAAGCTGAGCTGCCCCGAGGAGGGAGGAGTTC

- TGGGCTTCTGGCACCTGTTTCTAGGCCTAACCATTAGTACTTACTGTGCAGGGAACCAACCAAGGCTCGAGAAATGCCGACAC
CCCGAGCGAGACCCCCAAGTGCACAAAGCTGAGTAAAAAGCTGCCCCCTTCAAACAGAACTAGACTCAGTTTCAATTCCATC
CTAAACCTCCTTTTAACCAAGCTTAGCTTCTCAAAGGCCTAACCAAGCCTTGGCACCGCCAGATCCTTTCTGTAGGCTAATTCC
TCTTGCCCAACGGCATATGGAGTGTCTTATTGCTAAAAAGGATTCGGTCTCCTTCAAAGAAGTTTATTTTTTGGTCCAGAGTA
5 CTTGTTTTCCCGATGTGTCCAGCCAGCTCCGCAGCAGCTTTTCAAGATGCACATATGCCCTGATTGCTGATCGTGTTTAACTTTT
TCTTTTCTGTTTTTATTTTTGGTATTAAGTCGTGCTTTATTTGTAAAAGCTGTTATAAATATATATATATAAATATATATAA
AAGGAAAATGTTTCAGATGTTTTATTTGTATAATTAATTGATTACACAGTGAGAAAAATGAATGTATTCCTGTTTTTGAAGAG
AAGAATAATTTTTTTTCTCTAGGGAGAGGTACAGTGTTATATTTTGGAGCCTTCCTGAAGGTGTAATAATGTAATAATTTTTT
ATCTATGAGTAAATGTTAAGTAGTTGTTTTAAAAATACTTAATAAAAAATAATCTTTTCTCTGGAAG
- 10 373
MAKTYDYLKLLLLIGDSGVGKTCLLFRFSEDAFNITTFISTIGIDFKIRTIELDGKKIKLQIWDTAGQERFRITITAYYRGAMGI
MLVYDITNEKSFNINWIRNIEBHASSDVERMILGNKCDMNDKRQVSKERGEKLADYGIKFLETSAKSSANVEEAFTLARD
IMTKLNRKMNDNSAGAGGPVKITENRSKKTSPFRCSLL
- 374
15 ATTTGGCGGGCAGCCCCCTCGCCCGGGCCCCCTCCCCGCTCTCTCCACCGCCTCCTCTGGCTCCCCGGTCAGAGGGCCGGAG
CGAGAAGATGGCGAAGACGTACGATTATCTCTCAAGCTCCTGCTGATCGGCGACTCGGGGGTAGGCAAGACCTGCCTCCTGTT
CCGCTTCTCAGAGGACGCCCTTCAACACCACCTTCATCTCCACCATCGGAATTGATTTTTAAATTTAGAACGATAGAAGTACAGTGG
AAAGAAAATTAAGCTTCAGATATGGGACACAGCGGGTCAGGAAAGATTCGAACAATCACGACAGCGTACTACAGAGGAGCCAT
GGGCATTATGCTGGTCTATGACATCACAATGAAAAATCCTTTGACAATATTAAAAATTTGGATCAGAAACATTGAAGAGCATGC
20 CTCTCCGATGTCGAAAGAATGATCCTGGGTAACAAATGTGATATGAATGACAAAAGACAAGTGTCAAAAGAAAGAGGGGAGAA
GCTAGCAATGACTATGGGATTAAATTCCTGGAGACAAGCGCAAAATCCAGTGCAAAATGTAGAAGAGGCATTTTTTACACTTGC
ACGAGATATAATGACAAAACCTCAACAGAAAAATGAATGACAGCAATTCAGCAGGAGCAGGTGGACCAGTGAAAAATAACAGAAAA
CCGATCAAGAAGACCACTTCTTTTCGTGCTCGCTACTTTGATGAACCTTTCTGAGAGACTGCAGCACACCTAGAGGGCCCT
TTCTGCTTCTCTGAAAGCACAGGTACCCAGCCTCAGAATCACACCTCCCGGCTGCTGCTGAGAGCACCAGTGAACCTTAGACC
25 TCTCAACACAGTATGCCAAGTGGATTCCAGCCTCATGGCTAGCAAAAGAACAGACTCCCTTTTTCAACATGGAAGCAATGAA
GTGGAGACACATGCAGGACCTAATCGTTTTTTCCTTGTTTTATTACCTGTTGCAGAAGCGGTTATCTTTCTTTTACTTTGC
ACATCAGTGTAGCCTTTCCCTATTTTACGACAACTCTTAGACTCATATTTGCACACTTTTGTGTCGTGAAGTTCTAGACAAAT
TGATACATGTCGCAATGTTAAAAGAGCATTTACAGCAGAGGTTAATATACTAAAAATAAGGGTATTTGGTCTGGTTTCATATGGT
CAATATATTACTGCCCTTGGTAGCATTTATTTAAGGCTTTTCTTTAAATAAGAATCATTAAGTCATTAAAAAATTTACTGAAAT
30 GCCCA
- 375
MPGMFFSANPKELKGTTHSLDDKMQRPKTFGMDMKAYLRSMIPHLESGMKSSKSKDVL SAAEVMQWSQSLEKLLANQTGQN
VFGSFLKSEFSEENIEFWLACEDYKKTESDLLPCKAEIYKAFVHSDAAKQINIDFRTRESTAKKIKAPTPTCFDEAQKVIYTL
MEKDSYPRFLKSHIYLNLLNDLQANSLK
- 35 376
GGGAGTTAGACAAAATGCCAGGAATGTTCTTCTCTGCTAACCACAAAGGAATGAAAGGAACCACTCATTCACTTCTAGACGACA
AAATGCAAAAAAGGAGGCCAAAGACTTTTGAATGGATATGAAAGCATACCTGAGATCTATGATCCACATCTGGAATCTGGAA
TGAAATCTTCCAAGTCCAAGGATGTACTTTCTGCTGCTGAAGTAATGCAATGGTCTCAATCTCTGGAAAACTTCTTGCCAACC
AACTGGTCAAAATGTCTTTGGAAGTTTCTTAAAGTCTGAATTCAGTGAGGAGAAATATTGAGTTCTGGCTGGCTTGTGAAGACT
40 ATAAGAAAACAGAGTCTGATCTTTTGGCCCTGTAAAGCAGAAGAGATATATAAGCATTTGTGCATTTCAGATGCTGCTAAACAAA
TCAATATTGACTTCCGCACCTCGAGAATCTACAGCCAAGAAGATTAAAGCACCAACCCCCACGTGTTTTGATGAAGCACAAAAAG
TCATATATACTCTTATGGAAAAGGACTCTTATCCAGGTTCTCTCAAATCACATATTTACTTAAATCTCTTAAATGACCTGCAGG
CTAATAGCCTTAAAGTGACTGGTCCCTGGCTGAAGGGAATTAACAGATAGTATCAGCGCAGAAGGAATGTGCCAGTATGGATCCC
TGGGTGAACAGCTTGGCCTTTTTTGGGTGCTTGACAGGCCAAGAAGAACAATGACTCAGAACCAGGATTAACATGAAAGTTAT
45 CCAGGCGCAGAGTTGAAGAAGCATAGCAAGCAAGACAAAAACAGAGAGACCGCAAGGAGGAAGATCTGTGGTACTGTCTATAAA
AAACAGTGGAGCTCTGTATTAGAAAAGCCCCCTCAGAACTGGGAAGGCCAGGTAACCTTAGTTACACAGAACTGGTACTAAAGT
CTATCAAACTGATTACACAGACTGTAAGAATTCAAAGTCAACTGACATCTATGCTACATATATTATATAGTTTGTACTTGACTA
TGAGCCATTAACTTAAAGCATATGTTTCAATAGCCATTGCTACTATTCCTTGTCCGGTGAATTTTTATTTTATTTGTTTTTACT

TTGGAAGAGATGAACCTGTGTATTTAACTTAAGCTATTGCTCTTAAACCAGGGAGTCAGATATATTTGTAAGGTTAAATCATTG
GTGCAATAATAAATGTGGATTTTGTATTAAATATATATGAAGCAAAAAA

377

5 MQIFVKTLTGKTTITLEVEPSDTIENVKAKIQDKEGIPPDQRLIFAGKQLEDGRITLSDYNIQKESTLHLVLRRLRGGMQIFVKTL
TGKTTITLEVEPSDTIENVKAKIQDKEGIPPDQRLIFAGKQLEDGRITLSDYNIQKESTLHLVLRRLRGGMQIFVKTLTGKTTITLE
VEPSDTIENVKAKIQDKEGIPPDQRLIFAGKQLEDGRITLSDYNIQKESTLHLVLRRLRGGMQIFVKTLTGKTTITLEVEPSDTIE
NVKAKIQDKEGIPPDQRLIFAGKQLEDGRITLSDYNIQKESTLHLVLRRLRGGMQIFVKTLTGKTTITLEVEPSDTIENVKAKIQD
KEGIPPDQRLIFAGKQLEDGRITLSDYNIQKESTLHLVLRRLRGGMQIFVKTLTGKTTITLEVEPSDTIENVKAKIQDKEGIPPDQ
QRLIFAGKQLEDGRITLSDYNIQKESTLHLVLRRLRGGMQIFVKTLTGKTTITLEVEPSDTIENVKAKIQDKEGIPPDQRLIFAGK
10 QLEDGRITLSDYNIQKESTLHLVLRRLRGGMQIFVKTLTGKTTITLEVEPSDTIENVKAKIQDKEGIPPDQRLIFAGKQLEDGRITL
SDYNIQKESTLHLVLRRLRGGMQIFVKTLTGKTTITLEVEPSDTIENVKAKIQDKEGIPPDQRLIFAGKQLEDGRITLSDYNIQKE
STLHLVLRRLRGGV

378

15 TAGTTCCTCGCAGCCGGGATTTGGGTCGCGGTCTTGTGTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGA
AGACTCTGACTGGTAAGACCATCACCTCGAGGTTGAGCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG
AAGGCATCCCTCTGACCAGCAGAGGCTGATCTTTGCTGGAACACAGCTGGAAGATGGGCGCACCTGTCTGACTACAACATCC
AGAAAGAGTCCACCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAATCTTCGTGAAGACACTCACTGGCAAGACCATCA
CCCTTGAGGTGGAGCCAGTGACACCATCGAGAAGCTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCTCTGACCAGCAGAG
GGTTGATCTTTGCCGGAAGCAGCTGGAAGATGGGCGCACCTGTCTGACTACAACATCCAGAAAGAGTCTACCTGCACCTGG
20 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCTGACTGGTAAGACCATCACCTCGAGGTGGAGCCAGTGACA
CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCTCTGATCAGCAGAGGTGATCTTTGCCGGAAGCAGC
TGGAAGATGGTTCGTAACCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGACCTGGTACTCCGTCTCAGAGGTGGGATGC
AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCTTGAGGTGAGCCAGTGCACATTCGAGAAGCTCAAAGCAAAGA
TCCAAGACAAGGAAGGCATTCTCTGACCAGCAGAGGTTGATCTTTGCCGGAAGCAGCTGGAAGATGGGCGCACCTGTCTG
25 ACTACAACATCCAGAAAGAGTCTACCTTGACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCTGACTG
GTAAGACCATCACCTCGAAGTGGAGCCGAGTGACACCATTTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC
CTGACCAGCAGAGGTTGATCTTTGCCGGAAGCAGCTGGAAGATGGTTCGTGACTACAACATCCAGAAAGAGTCCA
CCTTGACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCTGACTGGTAAGACCATCACCTCGAGGTGG
AGCCGAGTGACACCATTTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCTGATCAGCAGAGGTTGATCTTTG
30 CTGGGAAACAGCTGGAAGATGGACGCACCTGTCTGACTACAACATCCAGAAAGAGTCCACCTGCACCTGGTCTCCGTCTTA
GAGGTGGGATGCAGATCTTCGTGAAGACCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTTGAGAATG
TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC
GCACCTGTCTGACTACAACATCCAGAAAGAGTCCACCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA
AGACCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTTGAGAATGTCAAGGCAAAGATCCAAGACAAGG
35 AAGGCATCCCTCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCTGTCTGACTACAACATCC
AGAAAGAGTCCACCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCTGACTGGTAAGACCATCA
CCCTCGAGGTGGAGCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCTGATCAGCAGAG
GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACCTGG
TCCTGCGCTTGAGGGGGGTGTCTAAGTTTCCCTTTTAAGGTTTCAACAAATTTTCATTGCACCTTCTCTTCAATAAAGTTGTT
40 GCATTCCC

379

MGSRTPEPLHAVQLRWGPRRRPPLVPLLLLLVPPPPRVGGFNLDAEAPVLSGPPGSFFGFSVEFYRPGTDGVSVLVGA
TSQPGVLQGGAVYLCWPWASPTQCTPIEFDSKSRLLLESSLSSEGEPEVEYKSLQWFGATVRAHGSSILACAPLYSWRTEKEP
LSDPVGTCYLDNFTRILEYAPCRSDFSWAAGQGYCQGGFSAEFTKTGRVVLGGPGSYFWQGGQILSATQEIQIESYYPYLLIN
45 LVQGGQLTRQASSIYDDSYLGSYVAVGEFSGDDTDFVAGVPKGNLTGYVTILNGSDIRSLYNFSGEQMASYFGYAVAATDVN
GDGLDLDLVGAPLLMDRTPDGRPQEVGRVYVYLQHPAGIEPTPLTLTGHDGFRFGSSLTPLGLDLDQDGYNDVAIGAPFGGET
QQGVVVFVPGGPGGLGSKPSQVLQPLWAASHTPDFFGSALRGGRLDGNYPDLIVGSFGVDKAVVYRGRPIVSASASLTIFPA
MFNPEERSCSLEGNPVACINLSFCLNASGKHVADSIGFTVELQLDWQKQKGGVRRALFLASRQATLTQTLLIQNGAREDCREMK
IYLRNESEFRDKLSPIHIALNFSLDPOAPVDHGLRPLHYQSKSRIEDKAQILLDCGEDNICVPDLQLEVFGEQNHVYLGDKN

ALNLTFFHAQNVGEGGAYEAEELRVTAPEAEYSGLVRHPGNFSSLSCDYFAVNQSRLLVCDLGNPMKAGASLWGLRFTVPHLRD
TKKTIQPDFQILSKNLNNSQSDVVSFRLSVEAQAQVTLNGVSKPEAVLFPVSDWHPRDQPQKEEDLGPVHVHVELINQGPSSI
SQGVLELSCPQALEGQQLLYVTRVTGLNCTTNHPINPKGLELDPEGSLHHQKREAPSRSSASSGPQILKCPAEACFRLRCELG
PLHQQESQSLQLHFRVWAKTFLQREHQPFSLQCEAVYKALKMPYRILPRQLPQKERQVATAVQWTKAEGSYGVPLWIIILAILF
5 GLLLLGLLIYILYKLGFFKRSLPYGTAMEKAQLKPPATSDA

380

CAGGACAGGAAGAGCGGGCGCTATGGGGAGCCGGACGCCAGAGTCCCCCTCTCCACGCCGTGCAGCTGCGCTGGGGCCCCCGGC
GCCGACCCCGCTCGTGCCGCTGCTGTTGCTGCTGCTGCCGCCGCCACCCAGGGTCGGGGGCTTCAACTTAGACGCGGAGGCC
CAGCAGTACTCTCGGGGCCCGGGCTCCTTCTCGGATTTCTCAGTGGAGTTTTACCGGCCGGAACAGACGGGGTCAGTGTGC
10 TGGTGGGAGCACCAAGGCTAATACCAGCCAGCAGGAGTGTGCGAGGGTGGTGTGTCTACCTCTGTCTTGGGGTGCCAGCC
CCACACAGTGCACCCCATTTGAATTTGACAGCAAAGGCTCTCGGCTCCTGGAGTCTCAGTGTCCAGCTCAGAGGGAGAGGAGC
CTGTGGAGTACAAGTCTTGCAGTGGTTTCGGGGCAACAGTTCGAGCCCATGGCTCCTCCATCTTGGCATGCGCTCCACTGTACA
GCTGGCGCACAGAGAAGGAGCCACTGAGCGACCCCGTGGGCACCTGCTACCTCTCCACAGATAACTTCAACCGAATTTCTGGAGT
ATGCACCCTCGCCGCTCAGATTTTCACTGCGGACAGGACAGGGTTACTGCCAAGGAGGCTTCACTGCGGAGTTCACCAAGACTG
15 GCCGTGTGGTTTTAGGTGGACAGGAAGCTATTTCTGGCAAGGCCAGATCTGTCTGCCACTCAGGAGCAGATTGCAGAATCTT
ATTACCCCGAGTACCTGATCAACCTGGTTTCAAGGGCAGCTGCAGACTCGCCAGGCCAGTTCCATCTATGATGACAGCTACCTAG
GATACTCTGTGGCTGTGTGGTGAATTCAGTGGTGTGACACAGAAGACTTTGTTGCTGGTGTGCCCAAAGGGAACCTCACTTACG
GCTATGTACCATCTCTAATGGCTCAGACATTCGATCCCTCTACAACCTTCTCAGGGGAACAGATGCCCTCTACTTTGGCTATG
CAGTGGCCGCCACAGACGTCATGGGGACGGGCTGGATGACTTGTGTGGTGGGGCACCCCTGCTCATGGATCGGACCCCTGACG
20 GCGGCCCTCAGGAGGTGGGACGGGTCTACGTCTACCTGCAGCACCCAGCCGGCATAGAGCCACGCCACCCCTTACCTCACTG
GCCATGATGAGTTTGGCCGATTTGGCAGCTCCTTGACCCCCCTGGGGGACCTGGACCAGGATGGCTACAATGATGTGGCCATCG
GGGCTCCCTTTGGTGGGGAGACCCAGCAGGGAGTAGTGTGTGTAATTTCTGGGGGCCAGGAGGGCTGGGCTCTAAGCCTTCCC
AGGTTCCTGCAGCCCTGTGGGCAGCCAGCCACACCCAGACTTCTTTGGCTCTGCCCTTCGAGGAGGCCGAGACCTGGATGGCA
ATGGATATCCTGATCTGATGTGTGGGTCTTTGGTGTGGACAAGGCTGTGGTATACAGGGGCCGCCCATCGTGTCCGCTAGTG
25 CCTCCCTCACCATCTTCCCGCCATGTTCAACCCAGAGGAGCGGAGCTGCAGCTTAGAGGGGAACCTTGGCCCTGCATCAACC
TTAGCTTCTGCCTCAATGCTTCTGGAAAACAGTTGCTGACTCCATTGGTTTACAGTGGAACTTCAGTGGAGTGGCAGAAGC
AGAAGGGAGGGGTACGGCGGCACCTGTTCTGGCCCTCCAGGCAGGCAACCTTGACCCAGACCCCTGCTCATCCAGAATGGGGCTC
GAGAGGATTCAGAGAGATGAAGATCTACCTCAGGAACGAGTCAAGATTTTCAGAGCAAACCTCTCGCCGATTCACATCGCTCTCA
ACTTCTCCTTGGACCCCAAGCCCAAGTGGACAGCCACGGCCTCAGGCCAGCCCTACATTATCAGAGCAAGAGCCGGATAGAGG
30 ACAAGGCTCAGATCTTGTGGACTGTGGAGAAGACAACATCTGTGTGCCTGACCTGCAGCTGGAAGTGTGGGGAGCAGAACC
ATGTGTACCTGGGTGACAAGAATGCCCTGAACCTCACTTCCATGCCCAGAAATGTGGGTGAGGGTGGCGCTATGAGGCTGAGC
TTCGGGTACCCGCCCTCCAGAGGCTGAGTACTCAGGACTCGTCAGACACCCAGGGAACCTTCCAGCCTGAGCTGTGACTACT
TTGCCGTGAACAGAGCCGCTGCTGGTGTGTGACCTGGGCAACCCATGAAGGCAGGAGCCAGTCTGTGGGTGGCCTTCGGT
TTACAGTCCCTCATCTCCGGGACACTAAGAAAACATCCAGTTTGAATTCAGATCCTCAGCAAGAATCTCAACAACCTCGCAAA
35 GCGAGCTGGTTTCTTTCGGCTCTCCGTGGAGGCTCAGGCCAGGTACCCCTGAACGGTGTCTCAAGCCTGAGGCAGTGTAT
TCCCAGTAAGCGACTGGCATCCCCGAGACAGCCTCAGAAGGAGGAGACCTGGGACCTGCTGTCCACCATGTCTATGAGCTCA
TCAACCAAGCCCCAGCTCCATTAGCCAGGGTGTGTGGAATCAGCTGTCCCCAGGCTCTGGAAGGTGAGCAGCTCCTATATG
TGACCAGAGTTACGGGACTCAACTGCACCACCAATCACCCCATTAACCCAAAGGGCTGGAGTTGGATCCCGAGGGTTCCCTGC
ACCACCAGCAAAAACGGGAAGCTCCAAGCCGAGCTCTGCTTCTCGGGACCTCAGATCCTGAAATGCCCGGAGGCTGAGTGT
40 TCAGGCTGCGCTGTGAGCTCGGGCCCTGCACCAACAAGAGAGCCAAAGTCTGCAGTTGCATTTCCGAGTCTGGGCCAAGACTT
TCTTGCAGCGGGAGCACACGCAATTTAGCCTGCAGTGTGAGGCTGTGTACAAAGCCCTGAAGATGCCCTACCGAATCTGCCTC
GGCAGCTGCCCCAAAAGAGCGTCAAGTGGCCACAGCTGTGCAATGGACCAAGGCAGAAGGCAGCTATGGCGTCCCACTGTGGA
TCATCATCCTAGCCATCCTGTGTTGGCCTCCTGCTCCTAGGTCTACTCATCTACATCCTCTACAAGCTTGGATTCTTCAAACGCT
CCCTCCCATATGGCACCAGCCATGGAAAAGCTCAGCTCAAGCCTCAGCCACCTCTGATGCTTGAATCTTCAATTTACAGACT
45 CCCATTTCTGAAGAACCAGTCCCCCACCCTCATTTACTGAAAAGGAGGGGTCTGGGTACTTCTTTGAAGGTGCTGACGGCCAG
GGAGAAGCTCCTCTCCCCAGCCAGAGACATACTTGAAGGGCCAGAGCCAGGGGGGTGAGGAGCTGGGGATCCCTCCCCCCCAT
GCACTGTGAAGGACCTTGTTTACACATACCCTCTTTCATGGATGGGGGAACCTCAGATCCAGGGACAGAGGCCAGCCTCCCTGA
AGCCTTTGCATTTTGGAGAGTTTCTGAAACAACGGAAAGATAACTAGGAAATCCATTACAGTTCTTTGGGCCAGACATGCC
ACAAGGACTTCTGTCCAGCTCCAACCTGCAAGATCTGTCTCAGCCTTGCCAGAGATCCAAAAGAGCCCCAGTAAGAACC
50 TGGAACTTGGGGAGTTAAGACCTGGCAGCTCTGGACAGCCCCACCTGGTGGGCCAACAAAGAACATAACTATGCATGGTGGC
CCAGGACAGCTCAGGACAGATGCCACAAGGATAGATGCTGGCCAGGGCCAGAGCCAGCTCCAAGGGAATCAGAACTCAA

- 1 TGGGGCCAGATCCAGCCTGGGGTCTGGAGTTGATCTGGAACCCAGACTCAGACATTGGCACCAATCCAGGCAGATCCAGGACTA
TATTTGGGCTGCTCCAGACCTGATCTGGAGGCCAGTTACCCCTGATTTAGGAGAAGCCAGGAATTTCCAGGACCTGAAGG
GGCCATGATGGCAACAGATCTGGAACCTCAGCCTGGCCAGACACAGGCCCTCCCTGTTCCTCCAGAGAAAGGGGAGCCCACTGTC
CTGGGCTGCAAGATTTGGGTTCTGCCCTGCCAGCTGCACTGATGCTGCCCTCATCTCTCTGCCCCAACCTTCCCTCACCTTGG
5 CACCAGACACCCAGGACTTATTTAAACTCTGTGCAAGTGAATAAATCTGACCCAGTGCCCCACTGACCAGAACTAGAAAAA
AAAA
- 381
- MDHTEGLPAEPPPAHAPSPGKFGERPFPKRLTREAMRNYLKERGDQTVLILHAKVAQKSYGNEKRFFCPPPCVYLMGSGWKKKK
EQMERDGCSEQESQPCAFIGIGNSDQEMQQLNLEGNKYCTAKTLYISDSDKRKHFIKSVKMFYGNSSDIDGVLKRIKVISKPS
10 KKKQSLKNADLCIASGTVKVALFNRLRSQTVSTRYLHVEGNFHASQQWGAFFIHLDDDESEGEFTVRDVYIHYGQCKLVC
SVTGMALPRLIIMKVDKHTALLDADDPVSQLHKCAFYLKDTMYLCLSQERIIQFQATPCPKPENKEMINDGASWTIISTDKA
EYTFYEGMGPVLAPVTPVPVVESELQLNGGDDVAMLELTGQNFTPNLRVWFGDVEAETMYRCGESMLCVVVDISAFREGWRWVRQ
PVQVPVTLVRNDGIIYSTSLTFTYTPPEGPRPHCSVAGAILPANSSQVPNPESNTNSEGSYTNASTNSTSVTSSTATVVS
- 382
- 15 ATCCCTCCGGTTTTCCTCAGTCTCCACGTACGTCCCTCAAAGCGCGTCTTAAACCCGGATAACCGGAGCGCTCCCCATGGAC
CACACGGAGGGCTTGCCTCGGAGGAGCCGCTGCGCATGCTCCATCGCTGGGAAATTTGGTGAGCGGCTCCACCTAAACGA
CTTACTAGGGAAGCTATGCGAAATTTATTTAAAGAGCGAGGGATCAAACAGTACTTATCTTCATGCAAAAGTTGCACAGAAG
TCATATGGAATGAAAAAGGTTTTCCTCCACCTCCTTGTGTATATCTTATGGGCAGCGGATGGAAGAAAAAAGAACAA
ATGGAACCGCATGTTGTTCTGAACAAGAGTCTCAACCGTGTGCATTTATTGGGATAGGAAATAGTGACCAAGAAATGCAGCAG
20 CTAACCTTGAAGGAAAGAACTATTCACAGCCAAAACATTGTATATATCTGACTCAGACAAGCGAAAGCACTTCATTTTTTCT
GTAAAGATGTTCTATGGCAACAGTGATGACATTGGTGTGTTCTCAGCAAGCGGATAAAAGTCATCTCCAAACCTTCCAAAAAG
AAGCAGTCATTGAAAAATGCTGACTTATGCATTGCCTCAGGAACAAAGGTGGCTCTGTTTAATCGACTACGATCCAGACAGTT
AGTACCAGATACTTGCATGTAGAAGGAGTAATTTTCATGCCAGTTCACAGCAGTGGGAGCCCTTTTTTATTCATCTCTTGGAT
GATGATGAATCAGAAGGAGAAGAAATTCACAGTCCGAGATGCTTACATCCATTATGGACAACATGCAAACTTGTGTGCTCAGTT
25 ACTGGCATGGCACTCCCAAGATTGATAAATTATGAAAGTTGATAAGCATAACCGCATTATTGGATGCAGATGATCTGTGTACAA
CTCCATAAATGTGCAATTTTACCTTAAGGATACAGAAAGATGTATTTGTGCTTTCTCAAGAAAGAATAATTCATTTACGGCC
ATCCCATGTCCAAAAGAACCAAAATAAGAGATGATAAATGATGGCGCTTCCTGGACAATCATTAGCACAGATAAGGCAGAGTAT
ACATTTTATGAGGGAATGGGCCCTGTCTTGGCCAGTCACTCCTGTGCTGTGGTAGAGAGCCTTCAGTTGAATGGCGGTGGG
GACGTAGCAATGCTTGAACCTACAGGACAGAATTTCACTCCAAATTTACAGTGTGGTTTGGGGATGTAGAAGCTGAAACTATG
30 TACAGGTGTGGAGAGAGTATGCTCTGTGTCGTCCAGACATTTCTGCATTCCGAGAAGGTTGGAGATGGGTCCGGCAACAGTC
CAGGTTCCAGTAACCTTGGTCCGAAATGATGGAATCATTTATTCACCAGCCTTACCTTTACCTACACACCAGAACAGGGCCA
CGGCCACATTGCAGTGTAGCAGGAGCAATCTTCAGCCAATTCAGCCAGGTGCCCTTAACGAATCAAACACAAACAGCGAG
GGAAGTTACACAAACGCCAGCACAAATTCACCAGTGTACATCATCTACAGCCACAGTGGTATCCTA
- 383
- 35 MESEPERKRARTDEVLPPEEAAPRRKMRTRTTCPMCRYAAPQLLLQKLLQRRRKGAEEEEQQDSGSEPRGDEDDIPLGPQSNV
SLLDQHQLKEKAEARKESAKEKQLKEEEKILESVAEGRALMSVKEMAKGITYDDPIKTSWTPPRYVLSMEERHERVRKYYHI
LVEGDGIPPIKSFEMKFPAAILRLKKGIHHTPIQIQIPTILSGRDMIGIAFTGSGKTLVFTLPVIMFCLQEKEKRLPFS
KREGPYGLIICPSRELARQTHGILEYYCRLLEDSSPLRLCALCIGMSVKEQMETIRHGVHMMVATPGRMLDLQKKMVSLDI
CRYLALDEADRMIDMGFECDIRTFISYFKGQRTLLFSATMPKKIQNFASALVKPVTINVGRAGAASLDVIEVEYVKEEAKM
40 VYLLECLQKTPPPVLIFAEKKADVDAIHEYLLKGVFAVHGGKDQEBRTKAIEAFREGKDVLVATDVASKGLDFPAIQHVI
NYDMPEEIEINYVHRIGRTGRSGNTGIATTFINKACDESVMIDLKALLLEAKQKVPVQLVLCGDESMLDIGGERGCAFCGGLG
HRTDCPKLEAMQTKQVSNIGRKDYLAHSSMDF
- 384
- 45 AATGGAGGAGTCGGAACCCGAACGGAAGCGGGCTCGCACCGACGAGGTGCTGCCGAGGAAGCCGCTCCGAGGCGGAAGATGAG
GACGACGAGGACTACGTGCCCTATGTGCCGTACGACGCGCGCAGCTACTGCTCCAGAAGCTGCTGCACGGAAGACGCAAGGG
AGCTGCCGAGGAAGAGCAGCAGGACAGCGGTAGTGAACCCGGGAGATGAGGACGACATCCCGCTAGGCCCTCAGTCCAACGT
CAGCCTCTGGATCAGCACCAGCACCTTAAAGAGAAGGCTGAAGCGCGCAAAGAGTCTGCCAAGGAGAAGCAGCTGAAGGAAGA
AGAGAAGATCTGGAGAGTGTGCCGAGGCGGAGCATGTGTCAGTGAAGGAGATGGCTAAGGGCATTACGTATGATGACCC

CATCAAAACCAGCTGGACTCCACCCCGTTATGTTCTGAGCATGTCTGAAGAGCGACATGAGCGCGTGCGGAAGAAATACCACAT
CCTGGTGGAGGGAGACGGTATCCACCACCCATCAAGAGCTTCAAGGAAATGAAGTTTCCTGCAGCCATCCTGAGAGGCTGAA
GAAGAAAGGCATTCAACCACCAACACCCATTTCAGATCCAGGGCATCCCCACCATTCTATCTGGCCGTGACATGATAGGCATCGC
TTTCACGGGTTACAGGCAAGACACTGGTGTTCAGTTGCCCGTCATCATGTTCTGCCTGGAACAAGAGAAGAGGTTACCCCTCTC
5 AAAGCGCGAGGGGCCCTATGGACTCATCATCTGCCCCTCGCGGGAGCTGGCCCGGCAGACCCATGGCATCCTGGAGTACTACTG
CCGCTGCTGCAGGAGGACAGCTCACCCTCTGCGCTGCGCCCTCTGCATTGGGGGCATGTCCGTGAAAGAGCAGATGGAGAC
CATCCGACACGGTGTACACATGATGGTGGCCACCCCGGGGCGCTCATGGATTTCGTGCAGAAGAAGATGGTCAGCCTAGACAT
CTGTCGCTACCTGGCCCTGGACGAGGCTGACCGCATGATCGACATGGGCTTCGAGGGTGACATCCGTACCATCTTCTCTACTT
CAAGGGCCAGCGACAGACCCTGCTCTTCAGTGCCACCATGCCGAAGAAGATTCAGAACTTTGCTAAGAGTGCCCTTGTAAGCC
10 TGTGACCATCAATGTGGGGCGCGCTGGGGCTGCCAGCCTGGATGTCATCCAGGAGGTAGAATATGTGAAGGAGGAGGCCAAGAT
GGTGTACCTGCTCGAGTGCCTGCAGAAGACACCCCGCTGTACTCATCTTTGCAGAGAAGAAGGCAGACGTGGACGCCATCCA
CGAGTACCTGCTCTCAAGGGGTTGAGGCGGTAGCCATCCATGGGGCAAGAGCAGGAGGAACGGACTAAGGCCATCGAGGC
ATTCCGGGAGGGCAAGAAGGATGTCTAGTACGACAGAGCTTGCCCTCCAAGGGCCTGGACTTCCCTGCCATCCAGCACGTCTAT
CAATTATGACATGCCAGAGGAGATTGAGAACTATGTACACGGATTGGCCGACCGGGCGCTCGGGAAACACAGGCATCGCCAC
15 TACCTTCATCAACAAAGCGTGTGATGAGTCAGTGCTGATGGACCTCAAAGCGCTGCTGCTAGAAGCCAAGCAGAAGGTGCCGCC
CGTCTGCGAGGTGCTGCATTGCGGGGATGAGTCCATGCTGGACATTGGAGGAGAGCGCGGCTGTGCCCTTCTGCGGGGCGCTGGG
TCATCGGATCACTGACTGCCCCAACTCGAGGCTATGCAGACCAAGCAGGTGAGCAACATCGGTGCAAGGACTACCTGGCCCA
CAGCTCCATGGACTTCTGAGCCGACAGTCTTCCCTTCTCTCAAAGAGGCTCAGTCCCCAAGACTGCCACCAGTCTACACATAC
AGCAGCCCCCTGGACAGAATCAGCATTTTCAGTCTAGCTGGCTGGAATGGGCCAGGCTGGTCTGGCTGCTGTTCCCTGTGCT
20 CTTGAGAATTACTGTTTTTGTTCCTTTTACCCAGCTGCCATTAAAGCCCAAACCTCTAGCCCAAAAAAAAAA

385

MKGDPNPKPRGMSSYAFFVQTCREEHKKKHPDSSVNFAEFSKKCSERWKTMSAKEKSKFEDMAKSDKARYDREMKNYVPPKGD
KKGKKKDPNAPKRPPSAFFLFCSEHRPKIKSEHPGLSIGDTAKKLGEMWSEQSAKDKQPYEQAAKLKEKEYEKDIAAYRAKGKS
EAGKKGPRPTGSKKKNEPEDEEEEEDEDEDEDEDEDE

25 386

GATGTGGCCCGTGGCCTAGCTCGTCAAGTTGCCGTGGCGCGGAGAACTCTGCAAAACAAGAGGCTGAGGATTGCGTTAGAGATA
AACCAGTTACGCGGAGCCCCGTGAGGGAAGCGTCTCCGTTGGGTCCGGCCGCTCTGCGGGACTCTGAGGAAAAGCTCGCACC
AGGTGGACCGGATCTGTCAACATGGGTAAAGGAGACCCCAACAGCCCGGGGCAAAATGTCTCGTACGCCCTTCTTCGTGCA
GACCTGCCGGGAAGAGCACAAGAAGAAACACCCGGACTCTTCCGTCAATTTCCGCGGAATTCTCCAAGAAGTGTTCCGGAGAGATG
30 GAAGACCATGTCTGCAAAAGGAGAAGTCGAAGTTTGAAGATATGGCAAAAAGTGACAAAGCTCGCTATGACAGGGAGATGAAAAA
TTACGTTCTCCCAAAGGTGATAAGAAGGGGAAGAAAAAGGACCCCAATGCTCCTAAAAGGCCACCATCTGCCCTTCTTCTGTT
TTGCTCTGAACATCGCCCCAAGATCAAAAGTGAACACCTTGCCCTATCCATTGGGGATACTGCAAGAAATTTGGGTGAAATGTG
GTCTGAGCAGTCAGCCAAAGATAAACAACCATATGAACAGAAAGCAGCTAAGCTAAAGGAGAAATATGAAAAGGATATGTCTGC
ATATCGTGCCAAAGGCAAAAGTGAAGCAGGAAAGAGGCCCCGTCAGGCCAACAGGCTCAAAGAAGAAGAACGAACCAGAAGA
35 TGAGGAGGAGGAGGAGGAAGAAGATGAAGATGAGGAGGAAGAGGATGAAGATGAAGAATAAATGGCTATCCTTTAATGATG
CGTGTGGAATGTGTGTGTGCTCAGGCAATTATTTTGCTAAGAATGTGAATTCAGTGCAGCTCAATACTAGCTTCAGTATAA
AACTGTACAGATTTTTGTATAGCTGATAAGATTCTCTGTAGAGAAATACTTTTAAAAAATGCAGGTTGTAGCTTTTTGATGG
GCTACTCATACAGTTAGATTTTACAGCTTCTGATGTTGAATGTTCTTAAATATTTAATGGTTTTTTAATTTCTGTGTGTATGG
TAGCACAGCAAACTTGTAGGAATTAGTATCAATAGTAAATTTTGGGTTTTTTAGGATGTTGCATTTTCGTTTTTTAAAAAAT
40 TTTGTAATAAAATATGTATATTATTTCTATTGCTTTTGTCTTAATATGCTAAGTTAATTTTCACTTTAAAAAAGCCATTTGAA
GACCAAAAAAAAAAAAAA

387

MRSRLLSVAHLPTIRETTEEMLLGGPGQEPSPSLDDYVRSISRLAQPTSVLDKATAQGQPRPPHRAQACRKRPAVSLRD
ITARFSGQOPTLPMDTVDFLDWLFGESQEQPSQRDLPRRTGPSAGLWGPHRQMDSSKPMGAPRGLCEARMPGHSRLRPPQD
45 GQQSSDLRSWTFGQSAQAMASRHRPRPSSVLRTLYSHLPVIEL

388

CACACTGCTCAGGGAAGAGCCTGCTACGGTGGACTGTGAGACTCAGTGCACTGTCTCTCTCCAGCGACCCACGCTGGACCCC
CTGCGGACCCCTCACCCCTCGGCCCCCAAGCTTCCAGGGGCTTCCTTTGGACTGGACTGTCCCTGCTCATCCATTCTCTGCTG

- CACCCCCAGACCTCCTCAGCTCCAGGTTGCCACCTCCTCTCGCCAGAGTGATGAGGTCCCGGCTTCTGCTCTCCGTGGCCCATC
TGCCCACAATTCCGGGAGACCACGAGGAGATGCTGCTTGGGGGTCTTGACAGGAGCCCCACCTCTCTAGCCTGGATGACT
ACGTGAGGTCTATATCTCGACTGGCACAGCCACCTCTGTGCTGGACAAGGCCACGGCCAGGGCCAACCCAGGCCACCCACA
GGCCAGCCAGGCTGCCGGAAGGGCCGCCCTGCTGTGTCCCTGCGAGACATCACCGCACGTTTCAGTGGCCAGCAGCCACAC
5 TGCCCATGGCTGATACTGTGGACCCCTGGACTGGCTTTTGGGGAGTCCCAGGAAAAGCAGCCAAGCCAGAGGGACCTGCCAA
GGAGGACTGGCCCTCTGCTGGCTCTGGGGTCCACATAGACAGATGGACAGCAGCAAGCCATGGGGGCCCCAGAGGGAGGC
TCTGTGAAGCCAGGATGCCCTGGGCATTCCCTGGCAAGACCACCGCAGGATGGGCAGCAGAGCTCTGACCTAAGAAGCTGGAATT
TTGGGCAGTCTGCCAAGCCATGGCCTCCCGCCACCGCCCCCGCCAGCAGTGTCTCAGAACCTCTACTCGCACCTCCCGG
TGATCCATGAACCTTGACCCCTCCCCAGTAAAGGCTTCTGTAGAGAGCATGCTGGGTCTGCATCTCCTCTCGTCTCCTCCATGG
10 TGGTCACTGCCCCCTGGCAGGTCTCTGAAAGGGAAATGCTTTTCTGCAGAGGCCCTTCTTGGGCAGTTACAGTTAGACCCACC
CCCTCTGAATATGATAACAGCCTGTTTACATGAGGAGATGTTACCAATCCCGTTTCGCTCTGACCCCTTGGCTGGCTGATCACCTT
GAGCAACTTACTTAACATCTGTGTTCCTCAGTTTCTCATGGGTAATATAGGGATAATTACTGGCACCTGCCTCCAGGCCATT
TGACGTGTACCGCATATAGGAGCCCACTGGCTGAGTAGCTACCATCATCGCTGGTGGGAAACCTGGTGGTAGGGGTGTGAGGGT
AGTGGGGGTGTTCAGCCCCAGGTGTTTCAAGAACAGGCCCTCGGGCACTCCCAAGTCTGCTCTTGGCTCCACCCCTCAAAGCCC
15 ATGTTCTGTGAGGCCCAAGAGAACACATGGAGTCTTAGCAAATGCACTAATGTATTCCGGGGGACTGTACCTGGCACCACTGG
GGCACTCTGCTGGCTACAACCTCATACGTCCTGTGGTGGCATTGGGAGAGTTCCTCCCATGATGAGGGCCAAGATAGAATCTGTAC
CACTCAGTGCTACCATCCCAACCCCTACACCACTTCCACACAGGGGCCCTCATGGCATGGTCAGGGTCCAGCTGTGGGTGAGAG
CAGGGCACTGTCCAGCTGTCCACTGGGGAAGTCAAGATGTCTTAAGGCCAGGCCAGGGCATCTGGAGTCTGAAGGACCTAGT
TCCTAGAGGCATCTGGCAGCAAGAAGGTGAGGCATCAGGGAACGGGAATCAGGCTGGGACTGATCAGAGGTGAAGGGACAGAGA
20 GAGGAGAGGAGGAAGATTGAGCTGGGGCAACAGCCAAGCTCACCTGGCAGGTCTCTGCCACCTCCTTCTCTGTGAGCTGTCACT
CTAGGTTAATCTCTTTTTTGTGGCTATTTTTAATTGCTTTGGATTGTGTTAAATGTTTTCTGTCTTCTGTAACTGTGTTTTCT
CTGGAGATAGAATGTAACCATATTAAGGAAAGGTTTCAGACAAGCAATTACCCAGTTTCTTTATCTATAAAATGGGGACA
TCAGCAATGTCTTTCACACCTGCAAGGGTGGGAATTGCGCATGTGAACCTTGAGCTGCATTTATGAGCACTGTAGACAAATG
TTGTATCTGTAC
- 25 389
- MQSLMQAPLLIALGLLLATPAQAHKKPSQLSSFSWDNCDGKDPVIRSLTLEPDPIVPGNVTL SVVGSTSVPLSSPLKVDL
VLEKEVAGLWIKIPCTDYIGSCTFEHFCVDLMDLIPTGEBCEPELRITYGLPCHCFKEGTYSLPKSEFVFPDLELPSWLTTGNY
RIESVLSSSGKRLGCIKIAASLKGI
- 390
- 30 CTCAGCTTCTTTGCGTAACCAATACTGGAAGGCATTTAAAGGACCTCTGCCGCTCAGACCTTGCAAGTAACTCCGCCCTGACC
CACCTTCCCGATGCAGTCCCTGATGCAGGCTCCCTCCTGATCGCCCTGGGCTTGCTTCTCGCGACCCCTGCGCAAGCCCAAC
TGAAAAGCCATCCAGCTCAGTAGCTTTTCTGGGATACTGTGATGAAGGGAAGGACCTGCGGTGATCAGAAGCCTGACTC
TGGAGCCTGACCCATCGTCGTTCCTGGAATGTGACCTCAGTGTCGTGGGCAGCACCAGTGTCCCCCTGAGTTCTCCTCTGA
AGGTGGATTAGTTTGGAGAAGGAGGTGGCTGGCTCTGGATCAAGATCCCATGCACAGACTACATTGGCAGCTGTACCTTTG
35 AACACTTCTGTGATGTGCTTGACATGTTAATTCCTACTGGGAGGCCCTGCCAGAGCCCTGCGTACCTATGGGCTTCTTGCC
ACTGTCCCITCAAAGAAGGAACCTACTACTGCCCAAGAGCGAATTCGTTGTGCTGACCTGGAGCTGCCAGTTGGCTACCA
CCGGGAACCTACCGCATAGAGAGCGTCCTGAGCAGAGTGGAAGCGTCTGGGCTGCATCAAGATCGCTGCCTCTCTAAAGGGCA
TATAGCATGGCATCTGCCACAGCAGAATGGAGCGGTGTGAGGAAGGTCCCTTTTCTCTGTTTTGTGTTTGCCAAGGCCAACT
CCCCTCTCTGCCCCCTTTAATCCCTTTCTACAGTGAGTCCACTACCTCACTGAAAATCATTTTGTACCATTACATTTTA
40 GGCTGGGGCAAGCAGCCCTGACCTAAGGGAGAATGAGTTGGACAGTTCTTGATAGCCCAGGGCATCTGCTGGGCTGACCAGTT
ACTCATCCCGTTAACATTTCTCTCTAAAGAGCCTCGTTCATTTCCAAGCAGTTAAGGAATGGGAACAGAGTGTTTTAGGACCT
GAAGAATCTTTATGACTCTCTCTCTTTCTCTTTTTTTTTTGTCTACTAAGTTAAAAGCGAAGTGAGAGTATTAAAGTTTGT
TTCTCTCCGGCCCCCTGTTACAATGAAGGGGCAAAAGTATTGCTCTTAGTCTATTCTCCCTTAACTTCTGTGACTAATTTT
TATTTCTCTTAGATTGCCCCAATTAACTACTAGGTTGAGTGTATCTCTGGAGAGGTAGGGTGTGTGGGGGAGGAATCCCTTGG
45 GGGAGATATTAGAGTGCTCTGTGTTTACAACTCAGGTACCCGAGGGCTAGCAAGAGACTTAAATGACTGATAAGAACCG
TGAGAAACATGTTGCTTCCAGGCTTGATTTTCGATTTTTCGCTTTTTTTTTTGGGACGGAGTCTCGCTTTGTACCAGGCTGGA
GTGCAGTGGTGCAATCTACCTCACTGCAACCTCCGCCCTCTGGGTCAAGCAATTTCTCTGCTCAGCCTCCCAAGTAGCTTG
GACTACAGGCCCTGCCACCACGCCCGCTAATTTGTGTATTTTAGTAGAGATGGGGTTTACCATGTTGGCCAGGATGGTCTC
GATCTCTTGACCTCGTGATCCGTCCACCTTGGCCTTGCAAGCGCTGGATTACAGGCATGAGCCACTACCCAGCCGATTTTT
50 CCTTTTGTATTAAAGATGCTATTACAATGTAAATATTTCTTACACAGAAAGTCACAGCACATGTGCCCATTGATACAAGGCTGC

- TGAGGCCCTGGTCTCCAGTTGGAAATATAATTAAGGGTGGCAGGGACTGGAGTCAGTTGGAGAGTGCATAGCCAGTCTGTGAAGA
CAACTGCCAGATACTGGCAATACTCCAGCCTGGTGACAGAGTGAGACTCTGTCTCAAAAAAAAAAGTTTCAATGTTTACTCCTAG
AGAAGCCAAAAATCCAGATTTGTATATGAAATCTTACCATTTTAAAAAGATTGGCAGCTAATTATTTTTTTAAAAAGCTGTGCAG
TGTGATGTGTCCCAACGGACTGGCTCATGGGTGGCCACGTCAACCTCTGATCTCAGACCGTGCATGCCCTGTCTCTTAAAG
5 ACAACTCTGTGGCACCCTTTCTCCCTCCACAGGGCCAAAGCCATAGTGTCCGGTCCCAAGGACAAGGCTCTTCCAGTGCTAGG
AGAGGTATGAGCAGCCTCTCACCCTGTGAGCTGTGGGGATCACAAGGCTGCCTGCCTCAGTCTTGGAGTCTGTGGGTGAATGA
GGCAGATGGGAAAGAGCCTCACCAGCAGCTGCTTTTGGAGCAGGGGTCCAAGGAAGAGAGGGTGGCCTCGACATCAAACTGCCCT
GGATTTTCTACCACCTGTACATCATAACAACTTCTGAAACACACACCAGCCCTGAGTCTTGGGCTCATTTGAAGCCTGGAA
TAGCAATAAATCTTTTAACTTGCAAAAAAAAAAAAAAAAAAAAA
- 10 391
MPAALVENSQVICEVWASNLEEMRKIREIVPSYSIAMDTEPGVVVRPIGEFRSSIDYQYQLLRNVDLLKIIQLGLTFTNE
KGEYPSGINTWQPNFKNLTEDMYSQDSIDLLANSGLQFQKHEEBIDTLHFAELLMTSGVVLCDNVKWLSPHSGYDFGYMVKL
LTD SRLPEEEHEFLHILNLFSPSIYDVKYLKSKCNLKGGLQEVADQLDLQIRGRHQAGSDSLTGMAFFRMKELFFEDSIDD
AKYCGRLYGLGTGVAQKQNEVDVSAQEKMSILAIINMQQ
- 15 392
GCACGAGGGGGTAGAGGAAAAGAGCTCCGGGCCAGGGCTGCCGTCCGCCCGCTCGGGGAGTCAGCCGCCAGCCGCCAGCT
CGTCAGCCGCCACCAGCTTCGCCGGCCCTGTCCGTCCCGTAAAGCGGCCCTGCCGTTACCGGAAAGAGGAGCGTAAGATGAAA
GAGTATCAGACCAACATTTCTCTGGCTTGCACTGTAAACTAGTTAGCTGAAGACGACTTCTCAGGTTCTTCAGGATGCCCTGC
AGCACTTGTGGAGAATAGCCAGGTTATCTGTGAAGTGTGGGCCAGTAATCTAGAAGAAGAGATGAGGAAGATCCGAGAAATCGT
20 GCCCAGTTACAGTTATATTTGCCATGGACACAGAAATTTCCAGGTGTTGTGGTGGCAGCAATTTGGTGAATTTCTGTAGTTCCATAGA
TTACCAATATCAGCTTCTGCCGTGCAATGTTGACCTTTTAAAAATTTATCCAGCTGGGCCCTTACATTACAAAATGAGAAGGGAGA
GTATCCTTCTGGAATCAATACTTTGGCAGTTCAATTTCAAATTTAACTTTACAGAGGACATGTACTCCCAGGATTCCATAGATCT
CCTTGCTAACTCAGGACTACAGTTTCAGAAGCATGAAGAGGAAGGGATTGACACACTGCACCTTTGCAGAGCTGCTTATGACATC
AGGAGTGGTTCTCTGTGACAATGTCAAATGGCTTTCAATTTATAGTGGCTATGATTTTGGCTATATGGTAAAGTTGCTTACAGA
25 TTCTCGTTTGCCAGAAGAGGAACATGAATTTACATATTTCTGAACCTTTTCTCCCCATCCATTTATGATGTGAAATACCTGAT
GAAGAGCTGCAAAAAATCTTAAGGGAGGCTTTCAGGAAGTTGCTGATCAGTTGGATTGTCAGAGGATTGGAAGGCAGCAGCAGGC
AGGCTCAGACTCACTGCTGACAGGAATGGCTTTCTTTAGGATGAAAGAGTTGTTTGTGAGGACAGCATTTGATGATGCCAAGTA
CTGTGGGGGGCTCTATGGCTTAGGCACAGGAGTGGCCAGAAAGCAGAATGAGGATGTGGACTCTGCCAGGAGAAGATGAGCAT
CCTGGCGATTATCAACAACATGCAGCAGTGATGGCGCCAGGCTCTGCAGGGTGGGCTGATCCCAGAGTGGTCTTACTGTGCT
30 GACTGTGTACTTATCTTCCCCAAGAGAAAATGCTTCTTTTGAGCACACTGTACCTACCCTGCAATTGAGCAGAAAGACTTTTG
TTTTACTGAAGACAAAAGATGTTTTATTTTAGACCCAGAAGAGAGGAGTTTGCTCTGAATTTGTAAATAAGTCTTCCCCATTC
CTCATACTCGAGCCTCTCTCTCTGGTTGCCCTCTGCCACCAGCATCCATGGCTCATTTGACACCTTTTAAATATCAGGACAA
GTCTGAAACAAAGTAGTAAATGTATATACTCTTACCTGTTGTCTATTCTTTTCTTTTAAATTTGTTGCTAATCTCTGATAAT
GAAGATTTCTACTCTGATTTCTCAGCTGAGCTGTGAGGGCTTCCAGGGAAAATGGAACAAAATGGTGTCTTATAGTAATGGGTTG
35 TAGATACTGAGTCTTCCTTTCTCTTCTGACCTTCTCGAGGACATTTGCTTTCTCTCACACTTTGTAGTCTCTCTTTACATAT
TACTATATGAAAATGAATGTCTGTGCTGAAATTTGAAGACCAGATAATGAACTGAAAAGCAAAACAATTTTACTGAATCTG
TCTACCTTCATTATGAGAACTCCAGAATGAGTGTGACCACCTGAAGCATCTTTTAAAGTCTGTGTTCCATTGTGCCATTTCAGGT
TTGCTGTACATATGCATCATCTGAAATCATTTGAAATTTTGTACAATAAAATATCCTGGATTGTATCCTGAAGGAACTAGT
AAGATCAGATTTTGGGTGATGCTGTGTTGATTTTCAGTAATGTGATTTTCAGATGGTCACTGAGATTCTCCACTTCTCTACTC
40 CATTTATTTCTCTACTTTTCTTCCAGCAACCTTGAAACGTGAGGGAGATGGATTAAATGTGAGTAACAGGAATGTGCTTTAAA
AAGCTAGAGTGGTTACATTTAATCAGGCAGTAAGATAATTTGGGTCTTTCAGTTGTTTTGGAGTAATATCCACAACCTGGGTA
GGAAGCTCAGGACTTTTTCTTTAAAGCTAGTCATTTCAAAAGCATATTTGATTTTTTTGAATGACTACAGTATGGACAATTTT
AAAAACCAAAACCCACTTTGGATTGGTGGAAAGTAAAACTGGTAACTCACTCAAGTGAATGAATGGTCTTCGATTTTAAAGCT
TATGGGAACTCAATTTGAAATGATTAGAAAATGTCAAGTATTATAAGCTGGTATTTAAGATGCTGTGAAATACTATTTATGTT
45 TTTAATTTTGTAATAAAGATTTCTTTTTTAAAAAAAAAAAAAAAAAAAAA
- 394
TCATCTTCTCTTTTATTTCTGACAACATTATGACATTAATAAAATCCAAGAGGTATTGGATTAGTCTTGTGTTGGTCTCTGGTG
ATTTAGATAAGGCAGAACTGAGCCCTCGGAATGTATTATCTCAAAGAGCTAGTAGCAGCTGCTATGCAAAGTTCTAAGCCCCG
AGTCAAATCTGGGCATCTCCACAGATGTGTGGTAGGGCAGGGCTCCAGCTCTGTGTGAAGAAGAACAGGAATGAGAACAGC

TTTATTTTGTAGCGCCAAGGTCCTACTATGCTCACACCTGTAATCCCAGCTTTGAGAAGAGGATCTGGCACCAGAAGGGTCTGG
TAGAAGAAAGTGGGAAAGGTGAAGGACAGGGATCCTTTCCCTTCTGCTACAACCTAGGGACAGAGGCATTGACTTCCGCAAGG
GGGCCAAAGAAGGTAGCAGCTACTTGGGGTTTGTGTGGGAGATGGCCGACCACTCCAAGGGAGGTCAAATGGCTTTTTCAGG
ACGTGGAATCTCGTCCAAAAAATGG

5 395

MKVSAALLCLLLIAATFIPQGLAQPDAINAPVTCCYNFTNRKISVQRLASYRRITSSKCPKEAVIFKTIVAKEICADPKQKWVQ
DSMDHLDKQTOTPKT

396

10 CTAACCCAGAAACATCCAATTCCTCAAACCTGAAGCTCGCACTCTCGCCTCCAGCATGAAAGTCTCTGCCGCCCTTCTGTGCCTGC
TGCTCATAGCAGCCACCTTCATTCCCAAGGGCTCGCTCAGCCAGATGCAATCAATGCCCCAGTCACCTGCTGTTATAACTTCA
CCAATAGGAAGATCTCAGTGCAGAGGCTCGCGAGCTATAGAAGAATCACCAGCAGCAAGTGTCCCAAAGAAGCTGTGATCTTCA
AGACCATTGTGGCCAAGGAGATCTGTGCTGACCCCAAGCAGAAGTGGGTTTCAGGATTCCATGGACCACCTGGACAAGCAAACCC
AAACTCCGAAGACTTGAACACTCACTCCACAACCAAGAATCTGCAGCTAACTTATTTTCCCTAGCTTTCCCCAGACACCTG
TTTTATTTTATTATAATGAATTTTGTGTTGTTGATGTGAAACATTAATGCCTTAAGTAATGTTAATTTCTTATTATAAGTTATTGATG
15 TTTTAAGTTTATCTTTTCATGGTACTAGTGTTTTTTAGATACAGAGACTTGGGGAATATGCTTTTCTCTTGAACCACAGTTCTA
CCCTGGGATGTTTTCAGGGTCTTTGCAAGAATCATTAATACAAAGAATTTTTTTTAACATTTCCAATGCATGCTAAAATATTA
TTGTGGAATGAATATTTTGTAACATTATACACCAATAAATATATTTTTGTACAAAAA

397

MDYLLMIFSLFLVACQGAETAVLGAELSAVGENGGEKPTSPPPWRLRRSKRSCSSLMDEKCVFYCHLDI IWNTPEHVVPYG
20 LGSPRSKRALENLLPTKATDRENCQCASQDKKKCNFCQAGKELRAEDIMEKDWNHKKGKDCSKLGKKCIYQQLVRGRKIRR
SSEHRLRTRSETMRNSVKSSEFHPKLGKPSRERYVTHNRAHW

398

25 GGAGCTGTTTACCCCCAC TCTAATAGGGGTTCAATATAAAAGACCGGCAGAGAGCTGTCCAAGTCAGACGCGCCTCTGCATCTG
CGCCAGGCGAACGGGTCC TGCGCCCTCTGCAGTCCAGCTCTCCACCACCGCGCGTGCCTCTGCAGACGCTCCGCTCGCTGCC
TTCTCTCTGGCAGGCGCTGCCTTTTCTCCCCGTTAAAGGGCAC TTGGGCTGAAGGATCGCTTTTGAGATCTGAGGAACCCGAG
CGCTTTGAGGGACCTGAAGCTGT TTTCTCTCGTTTCTCTTTGGGTTCAGTTTGAACGGGAGGTTTTTGATCCCTTTTTTTTCAGA
ATGGATTATTTGCTCATGATTTTCTCTCTGCTGT TTTGTGGCTTGCCAAGGAGCTCCAGAAACAGCAGTCTTAGCGCTGAGCTC
AGCGCGGTGGGTGAGAACGCGGGGAGAAACCCACTCCCAGTCCACCCTGGCGGCTCCGCCGCTCCAGCGCTGCTCCTGCTCG
TCCCTGATGGATAAAGAGTGTGTCTACTTCTGCCACCTGGACATCATTTGGGTCAACACTCCCGAGCACGTTGTTCCGTATGGA
30 CTTGGAAGCCCTAGGTCCAAGAGAGCCTTGGAGAATTTACTTCCCACAAAGGCAACAGACCGTGAGAATAGATGCCAATGTGCT
AGCCAAAAGACAAGAAGTGCTGGAATTTTGGCCAAGCAGGAAAAGAACTCAGGGCTGAAGACATTTATGGAGAAAGACTGGAAT
AATCATAGAAGAAAGAAAAGACTGTTCCAAGCTTGGGAAAAAGTGTATTTATCAGCAGTTAGTGAGAGGAAGAAAAATCAGAAGA
AGTTTCAGAGGAACACCTAAGACAAACCAGGTCCGAGACCATGAGAAACAGCGTCAAAATCATCTTTTCATGATCCCAAGCTGAAA
GGCAAGCCCTCCAGAGAGCGTTATGTGACCCACAACCCAGCACATTTGGTGACAGACTCTCGGGCGCTGTCTGAAGCCATAGCCTC
35 CAGGAGAGCCCTCTGGCCGACTTGCATCTCCACCCTGGCTGGGATCAGAGCAGGAGCATCTCTGCTGTTGTTCCCTGACTGGC
AAAGGACACAGCTCTCGTTCAAACAATCCAAGAAAGGTTTAAGGAGTTCGCCCAACCATCTTACATGGCTTTCATCAGTGGTA
ACTGCTTTGGTCTCTTCTTTTCATCTGGGGATGACAATGGACCTCTCAGCAGAAAACACACAGTCACATTCGAATTC

400

40 GTGTTGGAATCAATCAAATGTCATTTCAGGAAGCTTCTTGCTGGAATCCGAAGCACAGCTGTGTCTGTACCTTGCTCAGCAGCC
TGGGGCCTGGGTGTCTCCTTGTCATCCACTGGTCCATTCTGCTCTGATTTTTTTGTTCCTTTTTGGAGGTTCCACTTTGGG
TTTGGGCTTTGAAATATATAGGGCTACAAGTACTTGTCAGCTCCTTAATTTTAGCTTCAATCTCTTTTGACTTGACAACCTGGATC
CATGGTCAAACCTCGCTTGTTCTGCAGATTTAGCTTGTATTATTCATCCACTCCATTGCTTCATTTGTGCTTTTTTCTACCTTTGT
CATGTCAGCAGCATCCAAATGATCATACTGGTCCCCCTGTTTTGAAAGAGCCGATTATTTTCATATACTGTTGGAATCTGTTTC
CCTAGTCTCT

45 401

MAMVSEFLKQAWFIENEEQEYVQTVKSSKGGPGSAVSPYPTFNPSSDVAALHKAIMVKGVDIATIIDILTKRNNARQQIKAAY
LQETGKPLDETLKKALTGHLEEVLLALLKTPAQFDAELRAAMKGLGTDEDTLIEILASRTNKEIRDINRVYREELKRDLAKDI
TSDTSGDFRNALLSLAKGDRSEDFGVNEDLADSDARALYBAGERRKGTDVNVFNTILTTRSYPLRRVFPQKYTKYSKDHMNKVL
DLELKGDI EKCLTAIVKCATSKPAFFAEKLHQAMKGVGTRHKALIRIMVSRSEIDMNDIKAFYQKMYGISLQAILDETKGDYB
5 KILVALCGGN
402
AGTGTGAAATCTTCAGAGAAGAATTCTCTTTAGTTCTTTGCAAGAAGGTAGAGATAAAGACACTTTTTCAAAATGGCAATGG
TATCAGAATTCCTCAAGCAGGCCTGGTTTATGAAAATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCATCCAAAGGTGGTC
CCGGATCAGCGGTGAGCCCTATCTACCTTCAATCCATCCTCGGATGTCGCTGCCTTGCTAAGGCCATAATGGTTAAAGGTG
10 TGGATGAAGCAACCATCATGTACATTCTAACTAAGCGAAACAATGCACAGCGTCAACAGATCAAAGCAGCATATCTCCAGGAAA
CAGGAAAGCCCTGGATGAAACACTTAAGAAAGCCCTTACAGGTCACTTGAGGAGGTTGTTTTAGCTCTGCTAAAACTCCAG
CGCAATTTGATGCTGATGAACCTCGTGCTGCCATGAAGGCTTGAAGTGAAGATACTCTAATTGAGATTTTGGCATCAA
GAACTAACAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGGAAGTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACA
CATCTGGAGATTTTCGGAACGCTTTGCTTTCTCTTGCTAAGGGTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTTGGCTG
15 ATTCAGATGCCAGGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGACAGACGTAAACGTGTTCAATACCATCCTTACCACCA
GAAGCTATCCACAACCTTCGACAGAGTGTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTCTGGACCTGGAGT
TGAAAGGTGACATTGAGAAATGCCTCACAGCTATCGTGAAGTGCGCCACAAGCAAACAGCTTTCTTTGAGAGAAGCTTCATC
AAGCCATGAAGGTGTTGGAACCTGCCATAGGCATTGATCAGGATTATGTTTCCCGTCTGAAATTGACATGAATGATATCA
AAGCATCTATCAGAAGATGTATGGTATCTCCCTTTGCCAAGCCATCCTGGATGAAACCAAGGAGATTATGAGAAAATCCTGG
20 TGGCTCTTTGTGGAGGAACTAAACATTCCCTTGATGGTCTCAAGCTATGATCAGAAGACTTTAATTATATATTTTATCTCTAT
AAGCTTAAATAGGAAAGTTTCTTCAACAGGATTACAGTGTAGCTACCTACATGCTGAAAAATATAGCCTTTAAATCATTTTAT
ATTATAACTCTGTATAATAGAGATAAGTCCATTTTAAATAATGTTTCCCAAACCATAAACCTTATACAAGTTGTTCTAGT
AACAATACATGAGAAAGATGTCTATGTAGCTGAAAATAAATGACGTACAAGAC
403
25 MPLNVSFTNRNYDLDDYSVQPYFYCDEEENFYQQQQSELQPPAPSEDIWKKFELLPTPPLSPSRSSGLCSPSYVAVTPFSLRG
DNDGGGGSFSTADQLEMVTELLGGDMVNQSFICDPDETFIKNIIQDCMWSGFSAAAKLVSEKLASYQAARKDSGSPNPARGH
SVCSTSSLYQLDLASAASECIDPSVFPYPLNDSSSPKSCASQDSSAFSPSSDLSLSTESSPQGSPEPLVLHEETPPTSSDS
EEEQEDEEEDVVSVEKRQAPGKRSESGSPSAGGHSKPPHSPVLKRVHSTHQNHYAAPPSTRKDYPAKRVKLDVSVRLRQI
SNNRKCTSPRSSDTEENVKRRTHNVLERQRRNELKRSFFALRDQIPELENNEKAPKVVLKATAYILSVQAEQKLISEEDLL
30 RKRREQLKHKLEQLRNSCA
404
CTGCTCGCGCCGCCACCGCCGGGCCCCGGCCGCTCCCTGGCTCCCTCGCTCGAGAAAGGGCAGGGCTTCTCAGAGGCTTGG
CGGGAAGAAAGAACGGAGGGAGGGATCGCGCTGAGTATAAAGCCGGTTTTCGGGGCTTTATCTAACTCGCTGTAGTAATCCA
GCGAGAGGCAGAGGAGCGAGCGGGCGGCGGCTAGGGTGAAGAGCCGGCGAGCAGAGCTGCGCTGCGGGCTCCTGGGAAG
35 GGAGATCCGAGCGAATAGGGGGCTTCGCTCTGGCCAGCCCTCCCGCTTGATCCCCAGGCCAGCGGTCCGCAACCCCTGGCC
GCATCCACGAAACTTTGCCCATAGCAGCGGGCGGGCACTTTGCACTGGAACCTACAACACCCGAGCAAGGACGCGACTCTCCCG
ACGCGGGGAGGCTATTTCTGCCCATTTGGGGACACTTCCCGCGGCTGCCAGGACCCGCTTCTCTGAAAGGCTCTCTTGCAGCT
GCTTAGACGCTGGATTTTTCGGGTAGTGGAAACAGCAGCTCCCGCGACGATGCCCTCAACGTTAGCTTACCAACAGG
AAGTATGACCTGACTACGACTCGGTGCAGCCGATTTCTACTGCGACGAGGAGGAGAACTTCTACCAGCAGCAGCAGAGC
40 GAGCTGCAGCCCCCGCGCGCCAGCGAGGATATCTGGAAGAAATTCGAGCTGCTGCCACCCCGCCCTGTCCCCTAGCCGCGGC
TCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTACACACCTTCTCCCTTCGGGGAGACAACGACGCGGTGGCGGAGCTTCTCC
ACGCGCCGACGAGTGGAGATGGTGACCGAGCTGCTGGGAGGAGACATGGTGAACAGAGTTTATCTGCGACCCGAGCAGCAG
ACCTTCATCAAAAACATCATCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCCGCGCCAAGCTCGTCTCAGAGAAGCTGGCC
TCCTACCAGGCTGCGCGCAAAGACAGCGGCAGCCGAACCCCGCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTG
45 CAGGATCTGAGCGCCGCGCTCAGAGTGCATCGACCCCTCGGTGGTCTTCCCTACCTCTCAACGACAGCAGCTCGCCCAAG
TCCTGCGCTCGCAAGACTCCAGCGCTTCTCTCCGCTCCTCGATTTCTGCTCTCTCTCGACGGAGTCTCCCCGAGGGCAGC
CCGAGCCCTGCTGCTCCATGAGGAGACACCGCCACCACCAGCAGCGACTCTGAGGAGGAACAAGAAGATGAGGAAGAAATC
GATGTTGTTTCTGTGGAAGAGGAGGCTCCTGGCAAAAGGTGAGAGTCTGGATCACCTTCTGCTGGAGGCCACAGCAAACT

CCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACACATCAGCACAACTACGCAGCGCCTCCCTCCACTCGGAAGGAC
TATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTGAGAGTCTGAGACAGATCAGCAACAACCGAAAAATGCACCAGCCCCAGG
TCCTCGGACACCGAGGAGAATGTCAAGAGGCGAACACACAACGTCTTGGAGCGCCAGAGGAGGAACGAGCTAAAAACGGAGCTTT
TTTGCCCTGCGTGACCAGATCCCGGAGTTGAAAAAATGAAAAGGCCCCCAAGGTAGTTATCCTTAAAAAAGCCACAGCATAC
5 ATCCTGTCCGTCCAAGCAGAGGAGCAAAAGCTCATTTCTGAAGAGGACTTGTTCGGGAAACGACGAGAACAGTTGAAACACAAA
CTTGAACAGCTACGGAACTCTTGTGCGTAAGGAAAAGTAAGGAAAACGATTCTTCTAACAGAAATGTCTGAGCAATCACCTA
TGAACCTTGTTCAAATGCATGATCAAATGCAACCTCACAACTTGGCTGAGTCTTGAGACTGAAAGATTAGCCATAATGTAAA
CTGCCTCAAATTTGACTTTGGGCATAAAAGAACTTTTTATGCTTACCATCTTTTTTTTTTTTAAACAGATTGTATTTAAGA
ATTGTTTTTAAAAAATTTTAA

10 405

MGKNKLLHPSLVLLLLLVLLPTDASVSGKPQYMLVPSLLHTETTEKGCVLVSYLNETVTVSASLESVRGNRSLFTDLEAENDVL
HCVAFAPVKSSSNEEVMFLTVQVKGPTQEFKRRITVMVKNEDSLTVFQTDKSIYKPGQTVKFRVVSMDENFHLNELIPLVYIQ
DPKGNRIAQWQSFQLEGLKQFSPLSSEPFQGSYKVVVQKKSGRTEHPFTVEEFVLKFEVQVTVPKIITILEEEMNVSVCG
LYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSGLNSHGCFYQVVKVFLKRRKEYEMKLHTEAQIQEEGTVELT
15 GRQSSEITRTITKLSFVKVDSHFQGIPIFFGQVRLVDGKGVPIPNKVIPIRGNEANYSNATTDEHGLVQFSINTTNVMGTSLT
VRVNYKDRSPCYQWVSEHEEAAHTAYLVFSPSKSFVHLEPMSHELPCGHTQTVQAHYILNGGTLGLKLSFYLLIMAKGG
IVRTGTHGLLVKQEDMKGHFSISIPVKSADIAPVARLLIYAVLPTGDVIGDSAKYDVENCLANKVDLSPSPSQSLPASHAHLRVT
AAPQSVCALRAVDQSVLLMKPDAELSASSVYNLLPEKDLTGFPGLNDQDDEDCINRHNVIYINGITYTPVSSSTNEKDMYSFLED
MGLKAFTNSKIRKPKMCPQLQYEMHGPEGLRVGFYESDVMGRGHARLVHVEEPHTETVRKYFPETWIWDLVVVNSAGVAEYGV
20 TVPDTITIEWKAGAFCLSEBAGLGISSTASLRAFQPFVVELTMPYSVIRGEAFTLKATVLNLYLPKCIIRVSQVLEASPAFLAVPVE
KEQAPHICANGRQTVSWAVTPKSLGNVFTVSAEALSESQELCGTEVPVPEHGRKDTVIKPLLVEPEGLEKETTFNSLLCPSG
GEVSEELSLKLPNVVEESARASVSVLGDILGSAMQNTQNLQMPYCGCEQNMVLFAPNIYVLDYLNQTLTPEVKSIAIGYL
NTGYQRQLNKHVDGYSYTFGERYGRNQNTWLTAFVLKTFARAYIFIDEAHITQALIWLSQLKQDNGCFRSGSLNNAIK
GGVEDEVTLISAYITIALLEIPLTVTHPVVRNALFCLSAWKAQEGDHGSHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVK
25 KDNSVHWERPQKPAKPVGHFYEPQAPSAEVEMTSYVLLAYLTAQAPTSEDLTATNIVKWITKQONAQGGFSSTQDTPVALHA
LSKYGAATFTRTGKAAQVTIQSSGTFSSKFQVDNNRLLQVSLPELPGEYSMKVTGEGCVYLQTSKYNILPEKEEFPFALG
VQTLFPQTCDEPKAHTSFQISLSVSYTGSRSASNMAIVDKMVSQFIPLKPTVKMLERSNHVSRTEVSSNHVLIYLDKVSNTLS
LFFTVLQDVPVRDLKPAIVKVYDYETDEFAIAEYNAPCSKDLGNA

407

30 MKLCVTVLSLLMLVAAFCSPALSAPMGSDPPTACCFSYTARKLPRNFVVDYETSSLCSPAVVFQTKRSKQVCADPSETWVQE
YVYDLELN

408

AGCACAGGACACAGCTGGGTTCTGAAGCTTCTGAGTTCTGCAGCCTCACCTCTGAGAAAACCTCTTTTCCACCAATACCATGAA
GCTCTGCGTGACTGTCTCTCTCTCATGCTAGTAGCTGCCTTCTGCTCTCCAGCGCTCTCAGCACCAATGGGCTCAGACCC
35 TCCCACCGCTGTGCTTTTCTTACACCGCAGGAAGCTTCTCGCAACTTTGTGGTAGATTACTATGAGACCAGCAGCCTCTG
CTCCACGCCAGCTGTGGTATTCCAAACAAAAGAAGCAAGCAAGTCTGTGCTGATCCAGTGAGACCTGGGTCCAGGACTACGT
GTATGACCTGGAAGTGAAGTGAAGTGTCTCAGAGACAGGAAGCTTTCAGGGAAGGTCACCTGAGCCCGGATGCTTCTCCATGAGA
CACATCTCTCCATACTCAGGACTCCTCTCCGAGTTCCTGTCCCTTCTCTTAATTTAATCTTTTTTATGTGCCGTGTATTGT
ATTAGGTGTCATTTCCATTATTTATATTAGTTTAGCCAAAGGATAAGTGTCCCTATGGGGATGGTCCACTGTCACTGTTTCTC
40 TGCTGTTGCAAATACATGATAACACATTTGATTCTGTGTGTTTTCATAATAAACTTTAAAAATAAAATGCAGACAGTT

409

MESRGPLATSRLLLLLLLLLLRLHTRQGWALRPVLPQSAHDPPAVHLSNGPGQEPIAVMTFDLTKITKSSSFEVRTWDPEGVI
FYGDTPNPKDDWFMGLRDRPEIQLHNHWAQLTVGAGPRLLDGRWHQVEVKMEGDSVLEVDGEEVLRLRQVSGPLTSKRHPIM
RIALGGLFPASNRLPLVPALDGLRDRSWLDKQAEISASAPTSLSRCDVESNPGIFLPPGTQAEFNLRIQPQHAEPWAFSL
45 DLGLKQAAGSGHLALGTPENPSWLSLHLQDQKVLSGSGPGLDLPLVLGLPLQLKLSMRVLSQGSKMALALPPLGLAPL
LNLWAKPQGRFLGALPGEDSSTSFCLNGLWAQQQLDQALNRSHEIWITHSCPSPGNGTDASH

410

ATGGAGAGCAGAGGCCCCACTGGCTACCTCGCGCCTGCTGCTGTTGCTGCTGTTGCTACTACTGCGTACACCCGCCAGGGATGG
GCCCTGAGACCTGTTCTCCCCACCCAGAGTGGCCACGACCTCCGGCTGTCCACCTCAGCAATGGCCAGGACAAGAGCCTATC
GCTGTATGACCTTTGACCTCACCAGATCACAAAACCTCCTCCTCCTTTGAGGTTTGAACCTGGGACCCAGAGGGAGTGATT
TTTTATGGGGATACCAACCTAAGGATGACTGGTTTATGCTGGGACTTCGAGACGGCAGGCCTGAGATCCAATGCACAATCAC
5 TGGGCCCAGCTTACGGTGGGTGCTGGACCACGGCTGGATGATGGGAGATGGCACCAGGTGGAAGTCAAGATGGAGGGGGACTCT
GTGCTGCTGGAGTGGATGGGAGGAGGTGCTGCGCCTGAGACAGGTCTCTGGGCCCCTGACCAGCAACGCCATCCCATCATG
AGGATTGCGCTTGGGGGGCTGCTCTTCCCCGCTTCCAACCTTCGGTTGCGCTGGTTCTCTGCCCTGGATGGCTGCCGCGG
GATTCCTGGCTGGACAAACAGGCCGAGATCTCAGCATCTGCCCCACTAGCCTCAGAAGCTGTGATGTAGAATCAAATCCCGG
ATATTTCTCCCTCCAGGACTCAGGCAGAATTCAATCTCCGAGACATTCCCCAGCCTCATGCAGAGCCCTGGGCTTCTCTTTG
10 GACCTGGGACTCAAGCAGGCAGCAGGCTCAGGCCACCTCTGCTCTTGGGACACCAGAGAACCATCTTGGCTCAGTCTCCAC
CTCCAAGATCAAAGGTGGTGTCTTCTTGGGTCGGGGCAGGCTGGATCTGCCCTGGTCTTGGGACTCCCTCTTCAGCTG
AAGCTGAGTATGTCCAGGTGGTCTTGAAGCAAGGGTCGAAGATGAAGGCCCTTGGCCCTGCCCTCCCTTAGGCCCTGGCTCCCTC
CTTAACCTCTGGCCAAGCCTCAAGGGCTCTCTTCTTGGGGCTTTACCAGGAGAAGACTCTTCCACCTCTTTTGTCTGAAT
GGCCTTTGGGCACAAGGTGAGAGCTGGATGTGAGCAGGCCCTGAACAGAAGCCATGAGATCTGACTCACAGCTGCCCCAG
15 AGCCAGGCAATGGCACTGACGCTTCCCATTA.

411

MADKLTRIAIVNHDCKPKKCRQECKKSCPVVRMGKLCIEVTPQSKIWISETLCIGCGICIKKCPFGALSIVNLPNLEKETTT
HRYCANAFKLHRLPIPRPGEVLGLVGTNGIGKSAALKILAGKQKPNLGKYDDPPDWQEIILTYFRGSELQNYFTKILEDDLKAI
KPQYVARFLRLAKGTGVSILDRKDETKQAIVCQQLDLTHLKERNVEDLSGGELQRFACAVVCIQKADIFMFDEPSSYLDVKQR
20 LKAAITIRSLINPDRIYIIVVEHDLVLDYLSDFTCCLYGVPSAYGVVTPMPSVREGINIFLDGVVPTENLRFDRASLVFKVAET
ANEEEVKKCMYKYPGMKKMGFEELAIIVAGEFTDSEIMVLMGNGTGKTTFIRMLAGRLKPDGEGEVPVLNVSYKPKQKISPKS
TGSVRQLLHEKIRDAYTHPQFVTDVMPKPLQIENIIDQEVQTLSSGELQVRRLRLCLGKPADVYLIDEPSAYLDSEORLMAARVV
KRFILHAKKTAFVVEHDFIMATYLDRIIVFDGVPSKNTVANSPQTLLAGMKNFLSQLEITFRDRPNRYRPRINKLNSIKDVEQ
KKSNGYFFLDD

25 412

CCGTCCTGAGACACGCTGTGTGGCTGAAAAGTGAAGGCAAGAGCTCATTTGGCCTCTGTGCTCCCTCCGCAAGGGATCGTTT
CTCCAGAAGAGCTGGATATTCTTTGCCCCAGTTATGGCAGACAAGTTAACGAGAATTGCTATTGTCAACCATGACAAATGTAAA
CCTAAGAAATGTCGACAGGAATGCAAAAAGAGTTGTCTGTAGTTTGAATGGGAAAATTATGCATAGAGGTTACACCCAGAGC
30 AAAATAGCATGGATTTCCGAACTCTTTGTATTGGTTGTGGTATCTGTATTAAGAAATGCCCTTTGGCGCCTTATCAATTGTC
AATCTACCAAGCAACTTGGAAAAAGAAACCACATCGATATTGTGCCAATGCCTTCAAACTTCACAGGTTGCCTATCCCTCGT
CCAGGTGAAGTTTGGGATTAGTTGGAACATAATGGTATTGGAAGTCAGCTGCTTAAAAATTTTAGCAGGAAAAACAAAGCCA
AACCTTGGAAAGTACGATGATCCTCTGACTGGCAGGAGATTTGACTTATTCCGTGGATCTGAATTACAAAATTACTTTACA
AAGATTCTAGAAGATGACCTAAAAGCCATCATCAACCTCAATATGTAGCCAGATTCTTAAGGCTGGCAAAGGGGACAGTGGGA
TCTATTTTGGACCGAAAAGATGAAACAAAGACACAGGCAATTGTATGTCAGCAGCTTGATTTAACCACCTAAAAGAACGAAAT
35 GTTGAAGATCTTTTCAAGAGGAGAGTTGCAGAGATTGCTTGTGCTGTCGTTTGCAACAGAAAGCTGATATTTTCATGTTTGAT
GAGCCTTCTAGTTACCTAGATGTCAAGCAGCGTTTAAAGGCTGCTATTACTATACGATCTCTAATAAATCCAGATAGATATATC
ATTGTGGTGAACATGATCTAAGTGTATTAGACTATCTCTCCGACTTCACTGCTGTTTATATGGTGTACCAAGCGCCTATGGA
GTTGTCACTATGCCCTTTTACTGTAAAGAGAAGGCATAAACATTTTGTGGATGGCTATGTTCCAACAGAAAACCTTGAGATTGAGA
GATGCATCACTTGTTTTAAAGTGGCTGAGACAGCAATGAAGAAGAAGTTAAAAGATGTGTATGTATAAATATCCAGGAATG
40 AAGAAAAAATGGGAGAAATTTAGCTAGCAATTGTAGCTGGAGAGTTTACAGATTCTGAAATTATGGTGTGCTGGGGGAAAT
GGAACGGGTAAAACGACATTTATCAGAATGCTTGTGGAAGACTTAAACCTGATGAAGGAGGAGAAGTACCAGTTCTAAATGTC
AGTTATAAGCCACAGAAAATTAGTCCCAATCAACTGGAAGTGTTCGCCAGTTACTACATGAAAAGATAAGAGATGCTTATACT
CACCACAATTTGTGACCGATGTAATGAAGCCTCTGCAAAATTGAAAACATCATTTGATCAAGAGGTGCAGACATTATCTGGTGGT
GAACTACAGCGAGTACGTTTACGCTTTGCTTGGGCAAACTGCTGATGCTATTTAATTGATGAACCATCTGCATATTGGAT
45 TCTGAGCAAAGACTGATGGCAGCTCGAGTTGTCAAACGTTTCATACTCCATGCAAAAAAGACAGCCTTTGTTGTGAACATGAC
TTCATCATGGCCACCTATCTAGCGGATCGCGTCATCGTTTGTGATGGTGTTCATCTAAGAACACAGTTGCAACAGTCTCAA
ACCTTTTGGCTGGCATGAATAAATTTTGTCTCAGCTTGAATTACATTCAGAAGAGATCCAAACAATATAGGCCACGAATA
AACAACTTAATTCAATTAAGGATGTAGAACAAAAGAGTGGAACTACTTTTCTTGGATGATTAGACTACTCTGAGAAT
ATTGATAAGCCATTTATTAAGAGGATTTTACTAGAATTTTGTGTCATATAAACTTGAATCAGGATTTTATGCCCCACATAC

- TCTGGAACCTTGAAGTATAATATACTTAATATAACATAAAAAGCCAGTTGGGTTCTAAATTTAGTGTGAAACACAGAAAAATGCCA
CTTTTCTGTTCCCTGAAGAGGCTCTTTTGTGCATAATATTTCTAAAATGAAGACATTTCAAGCTATACAAATTACTTCCAGGTTTTT
CATGATGTATGGGAAGATTTTCAGTAGGTGTATATATTCACGGTACCAAATGCTGACCAGTGTGCTCCATTTTAAATCTTT
GAAAAGGGTTCTGTACTTACCTGGTTTGCCAAGTATGCCAGTGAATGAAACTGCCCTTATTTTAAAGCCAGTCAAAGATTC
5 CACTGATTGACATTTGATAAATAAACATCAGGATTATGTTTATTTGTTTTCAGTCTTTGCACTATATTACCAGTATATGGT
TTCCGAGGAAGATTATCTACTGCAAAACACCAGTGTGGAAAAATAGGTATTTTAAATTTGTTTAAATCTTTTGGTGGCTT
TTAAACATGTTTAAAGCAAAAACCAATTCAGTCCATTTCCCGCAAAAACCCCTAACTTTACTCTGAACCTTTTGTGTTTGGCA
TTCCATGAGGTTCGTATTCAGTCATTTCTCTAGGTAATGTCAATTTTGTACACATATATTTATATAATCACTGATTGAGATTTA
GGAAAAAGCATTTCTAAAGAATATTTGCTTCCCTTAGAACTACAGACTCGAAATCTTTAAAGATGGTGCCTAAGCATCTATGTA
10 TTTTCTTTTAAAGTCCACAGATTTTCTGTGTGGCAGGCCAAGGATTATAAACCACTTCCCTAAAGGCAACATTAATGCAAAAGT
CCCCAGATGGCAATAACAAAGTATCCCCGGTACCACATATATTCATTTGTGAGTTTGGATATAGAGCACATTATCTAAACCAAT
TTGTAGTTCCAAAAACCATCTAAATTTCTTGAGTTCTGAAATTTTGAACAGGATTACCTGGAGCCCTGGAGCCACTTTAAGTTG
TACTTCTGACTAACTGGAATTATGAGTGAGGAAGAGTGTTTACTAAATAAATGACTGGGGCAAGCAAAATTGAGGAGGAAAT
AGAACTGTTTGACAACTTTAAGAGCTACTTGAAATAACAGAAGTCTTGATTAAATGCAAAATAATGGCTAGAAAGTATGGTT
15 TAACTGGACCTTATATGCCTTTTAAAAATAATTCAGTAACCCATAAATACATGTTGTAAAAAATTCAAATATACAGAATGGA
ATAAAAAATGATCTCCCTTTATTACCTCCCAAAGGTACCAGCGTTGAATTTAATAATGTATATTTCTTTCATGCTTTTTTC
TGTGCACTTACCTAAGTGTGAATATGTAAAGGGTTGTTTGTATACAAATGGGATTATACTAAAATAAGTAATGCCTATTTTT
AAGGATAGGTTAAATTTGTGAATGATCATTTCAAATATATTGAATAAAATAAGCAAAAGCTATTGTTATTTACTGATCTGAAA
AAA
- 20 413
- MAAVKTLNPKAEVARAQAALAVNISAARGLQDVLRTNLGPKGTMKMLVSGAGDIKLTGDNVLLHEMQIQHPTASLIAKVATAQ
DDITGDGTTSNVLIIGELLKQADLYISEGLHPRIITEGFEEAKEKALQFLEEVKVSREMDRETLIDVARTSLRKVHAELADVL
TEAVVDSILAIIKKQDEPIDLFMIEIMEMKHKSETDTSLIRGLVDHGHARHPDMKKRVEDAYILTCNVSLYEYKTEVNSGFFYKS
AEEREKLVKAERKFIEDRVKKIIELKRKVCSDKGFVVINQKIDPFLDALSKEGIVALRRAKRRNMRERLTLACGGVALNSF
25 DDLSPDCIGHAGLVVEYTLGEEKFTFIEKCNNPRSVTLIKGNPKHNTLTQIKDAVRDGLRAVKNAIDGCVVPGAGAVEVAMAE
ALIKHKPSVKRAQLGVQAFADALLIIPKVLQNSGFDLQETLVKIQAEHSESGQLVGVDLNTGPEPMVAEVEGVWDNYCVKKQL
LHSCVTIATNILVDEIMRAGMSSSLKG
- 414
- CGAGAAGACCCGGATAGTTCTCCCGGCCACGCCGCGCGGCTCTGGGCACTCAGCATCGTTTCTCTTCTCCGCTGGAGCAG
30 CTATGGCGGCGGTGAAGACCTGAACCCCAAGGCCGAGGTGGCCGAGCGCAGGCGCGGCTGGCGGTCAACATCAGCGCAGCGC
GGGGTCTGCAGGACGTGCTAAGGACCAACCTGGGGCCCAAGGCCACCATGAAGATGCTCGTTTCTGGCGCTGGAGACATCAAAC
TTACTAAAGACGGCAATGTGCTGCTTACGAAATGCAAAATCAACACCCCAACAGCTTCTTAAAGCAAAAGGTAGCAACAGCCC
AGGATGATATACTGGTATGCTACGACTTCTAATGTCTAATCATTGGAGAGCTGCTGAAACAGGCGGATCTCTACATTTCTG
AAGGCCTTCATCTAGAAATAATCACTGAAGGATTTGAAGCTGCAAGGAAAAGGCCCTTCAGTTTGTGAAGAAGTCAAAGTAA
35 GCAGAGAGATGGACAGGAAACACTTATAGATGTGGCCAGAATCTCTTCTGTAAGTTTCAATGCTGAACTTGCAGATGTCT
TAACAGAGGCTGTAGTGGACTCCATTTTGCCCATTAAGCAAGATGAACCTATTGATCTCTTCATGATTGAGATCATGGAGA
TGAAACATAAATCTGAAACTGATACAAGCTTAATCAGAGGGCTTGTTTTGGACCACGGAGCACGGCATCCTGATATGAAGAAAA
GGGTGGAGGATGCATACATCTCACTTGTAACGTGTCTATTAGATGATGAGAAAAAGAGTGAATTCGGCTTTTTTTACAAGA
GTGCAGAAGAGAGAGAAAACTCGTGAAGCTGAAAGAAAATTCATTGAAGATAGGGTTAAAAAATAATAGAACTGAAAGGA
40 AAGTCTGTGGCGATTAGATAAAGGATTTGTTGTTTATTAATCAAAAGGGAATTGACCCCTTTTCTTATAGTGTCTTTCAAAAG
AAGGCATAGTCTGCTGCGCAGAGCTAAAAGGAGAAATATGGAGAGGCTGACTCTTGCTTGTGGTGGGTAGCCCTGAATCTCT
TTGACGACCTAAGTCTGACTGCTTGGGACATGCAGGACTTGTATATGAGTATACATTGGGAGAGAGAAAGTTTACCTTTATTG
AGAAATGTAAACCCCTGTTCTGTCTACATTATTGATCAAGGACCAATAAGCACACACTCACTCAGATCAAGATGCAGTGA
GGGACGGCTTGAGGGCTGTCAAAAATGCTATTGATGATGGCTGTGTGGTTCCAGGTGCTGGTGGCGTGAAGTGGCAATGGCAG
45 AAGCCCTGATTAACATAAGCCAGTGTAAAGGCGAGGCGACAGCTTGAGTCCAAGCATTGCTGATGCAATTGCTCATTTATT
CCAAGGTCTTGCTCAGAACTCTGGTTTTGACCTTCAGGAAACATTAGTTAAATTCAGCAGAAACATTGCAATCAGGTCAGC
TTGTGGGTGTGGACCTGAACACAGGTGAGCCAATGGTGGCAGCAGAAGTAGGCGTATGGGATAACTATTGTGTAAGAAACAGC
TTCTTCACTCTGCACTGTGATTGCCACCAACATTCTCTTGGTTGATGAGATCATGCGAGCTGGAATGTCTTCTCTGAAAGGTT
GAATTGAAGCTTCTCTGTATCTGAATCTTGAAGACTGCAAGTGATCTTGAGGATTACAGCTGTGGAATTTTGTCCAAGCTT
50 CAAATAATTTTGAAGAAATTTTCCCATATGAAAAAGGAGAGAACACTGGCATCTGTGAAATTTGGAAGTTCTGAAATTATA

5 MTAPGAAGRCPPPTTWLGSLLLLVCLLASRSITEEVSEYCSHMI GSGHLQSLQRLIDSQMETSCQITFEFVDQEQ LKDPVCYLKK
AFLLVQDIMEDTMRFRDNTANPIAIVQLQELSLRLKSCFTKDYEEHKACVRTFYETPLQLLEKVKNVFNETKNLLDKDWNIFS
KNCNNSFAECSSQDVVTKPDCNCLYPKAIPSSDPASVSPHQPLAPMAPVAGLTWEDSEGTEGSSLLPGEQPLHTVDPGSAKQR
PPRSTCQSFEPPETPVVKDSTIGGSPQRPSPVGA FNPGMEDILDSAMGTNWVPEEASGEASEIPVPQGTELSPPRPGGSGMQTE
PARPSNFLSASSPLPASAKGQQPADVTATALPRVGPVMPGTGQDWNHTPQKTDHPSALLRDPPEGPSRRISSLRPQALSNPSTLS
AQPQLSRSHSSGSVLP LGELEGRSTRDRTPSAEPEAAPASEGAARPLPRFNSVPLTDTGHERQSEGSSSPQLQESV FHLLVPS
10 VILVLLAVGGLLFYRWRRRSHQEPQRADSPLEQPEGSPLTQDDRQVELPV

CTGGGTCCTCTCGCGCCAGAGCCGCTCTCCGCATCCAGGACAGCGGTGCGGCCCTCGGCCGGGGCCCACTCCGCGACGAG
CCAGCGAGCCAGCTGCCCCGTATGACCGCGCCGGGCGCCGCGGGCGCTGCCCTCCCACGACATGGCTGGGCTCCCTGCTGTTG
TTGGTCTGTCTCCTGGCGAGCAGGAGTATCACCGAGGAGGTGTGCGAGTACTGTAGCCACATGATTGGGAGTGGACACCTGCAG
TCTCTGCAGCGGCTGATTGACAGTCAGATGGAGACCTCGTGCCAAATTACATTTGAGTTTGTAGACCAGGAACAGTTGAAAGAT
CCAGTGTGCTACCTTAAGAAGGCATTTCTCCTGGTACAAGACATAATGGAGGACACCATGCGCTTCAGAGATAACACCGCCAAT
CCCATCGCCATTGTGCAGCTGCAGGAACCTCTCTTTGAGGCTGAAGAGCTGCTTACCAAGGATTATGAAGAGCATGACAAGGCC
TGCCTCCGAACCTTCTATGAGACACCTCTCCAGTTGCTGGAGAAGGTCAAGAAATGCTCTTAATGAAACAAAGAAATCTCCTTGAC
AAGGACTGGAATATTTTCAGCAAGAACTGCAACAACAGCTTTGCTGAATGCTCCAGCCAAGATGTGGTGACCAAGCCGTGATTGC
AATCTGCCGTGACCCCAAAGCCATCCCTAGCAGTGACCCGGCCTCTGTCTCCCCCTCATCAGCCCTTCGCCCCCTCCATGGCCCCCT
GTGGCTGGCTTGACCTGGGAGGACTCTGAGGGAAGTGAAGGACAGCTCCCTCTTGCTGGTGGTGAGCAGCCCTGCACACAGTGGAT
CCAGGCAGTGCCAAGCAGCGGCCACCCAGGAGCACCTGCCAGAGCTTTGAGCCGCCAGAGACCCAGTTGTCAAGGACAGCACC
ATCGGTGGCTCACCACAGCCTCGCCCCCTCTGTCTGGGGCCTTCAACCCCGGGATGGAGGATATTTCTTGACTCTGCAATGGGCAT
AATTGGGTCCCAGAAGAAGCCTCTGGAGAGGCCAGTGAGATTCCCGTACCCCAAGGGACAGAGCTTTCCCCCTCCAGGCCAGGA
GGGGGCAGCATGCAGACAGAGCCCGCCAGACCCAGCAACTTCTCTCAGCATCTTCTCCACTCCCTGCATCAGCAAAGGGCCAA
CAGCCCGCAGATGTAACGTGTACAGCCTTGCCCAAGGTGGGCCCGGTGATGCCACTGGCCAGGACTGGAATCACACCCCCAG
AAGACAGACCATTCATCTGCCCTGTCTCAGAGACCCCCCGAGCAGGCTCTCCAGGATCTCATCACTGCGCCGCCAGGCCCTC
AGCAACCCCTCCACCCTCTGTCTGTCTCAGCCACAGCTTTCCAGAAGCCACTCTCGGGCAGCGTGTCTGCCCTTGGGGAGCTGGAG
GGCAGGAGGAGCACCAGGATCGGACGAGCCCCGACGAGCAGAGCAGCACCAGCAAGTGAAGGGGACGCCAGGCCCTGCCCC
CGTTTTAACTCCGTTCTCTTTGACTGACACAGGCCATGAGAGGCAGTCCGAGGGATCCTCCAGCCCGCAGCTCCAGGAGTCTGT
TTCACCTGTCTGGTGCCAGTGTATCTGGTCTTGTCTGGCTGTCTGGAGGCCCTTGTGTCTACAGGTGGAGGCGCGGAGCCAT
CAAGAGCCTCAGAGAGCGGATTCTCCCTTGGAGCAACCAGAGGGCAGCCCCCTGACTCAGGATGACAGACAGGTGGAAC TGCCA
GTGTAGAGGGAATTTAAGCTGGACGCACAGAACAGTCTCTTCTGTGGGAGGAGACATTATGGGGCGTCCACCACCACCCCTCCC
TGGCCATCTCTCGAATGTGGTCTGCCCTCCACCAGAGCTCTGCTGCCAGGACTGGACCAGAGCAGCCAGGCTGGGGCCCC
TCTGTCTCAACCCGACAGCCCTTGACTGAATGAGAGAGGCCAGAGGATGCTCCCCATGCTGCCACTATTTATTGTGAGCCCTGG
AGGCTCCCATGTGCTTGAGGAAGGCTGGTGAGCCCGGCTCAGGACCCCTCTTCCCTCAGGGGCTGCAGCCTCCTCTCACTCCCTT
CCATGCCGGAACCCAGGCCAGGGACCCACCGCCTGTGGTTTGTGGGAAAGCAGGCTGCACGCTGAGGAGTGAACAACCCCTGC
ACCCAGAGGGCTTGCTGTGTGCCAAGGTA TCCAGCCTTGACAGGCATGACCTGTCTCCAGACAGAGGAGCCCTGAAGTTCTGTG
GGGCGGGACAGCCTCGGCCGTGATTTCCCGTAAAGGTGTGCAGCCTGAGAGACGGGAAGAGGAGGCCTCTGCACCTGTCTGGTCTG
CACTGACAGCCTGAAGGGTCTACACCCCTCGGCTCACCTAAGTCCCTGTGCTGGTTGCCAGGCCCAGAGGGGAGGCCAGCCCTGC
CCTCAGGACCTGCCTGACCTGCCAGTGATGCCAAGAGGGGGATCAAGCACTGGSCCTTGCCCCCTCTCCTTCCAGCACCTGCCA
GAGCTTCTCCAGCAGGCCAAGCAGAGGCTCCCCCTCATGAAGGAAGCCATTGCACTGTGAACATGTACCTGCTGCTGAACAGC
CTCCCCCGTCCATCCATGAGCCAGCATCCGTCCGTCTCCACTCTCCAGCCTCTCCCCAGCCTCCTGCACTGAGCTGGCCTCA
CCAGTCTGACTGAGGGAGCCCCCTAGCCCTGACCTTCTCCTGACCTGGCCTTTGACTCCCCGAGTGGAGTGGGGTGGGAGAACC
TCCTGGGCCGCCAGCCAGAGCCGCTCTTTAGGCTGTGTTCTTCGCCCAGGTTTCTGCATCTTCCACTTTGACATTTCCCAAGAGG
GAAGGCACTAGTGGGAGAGACAAGGAGGGGAGGGCACAAGCAGAGAGCCTACAGGGCGAGCTGCACTGAAGATGGGCCTTT
GAAATATAGGTATGCACCTGAGTTGGGGGAGGGTCTGCATCCCAACCCACGCGAGTGTCTTTCCCTGTGCTGCCAGAGGA
ACCTGGGGCTGAGCAGGTTATCCCTGTTCAGGAGCCCTGGACTGGGCTGCATCTCAGCCCACTGCACATGTGTATCCAGCTCCCAT
CCACTTCTCACCCCTTCTTTCTCTCTGACCTTGGTTCAGCAGTGATGACCTCCAACCTCACCCACCCCTCTACCATCACCTCTA
ACCAGGCAAGCCAGGCTGGGAGAGCAATCAGGAGAGCCAGGCCTCAGCTTCCAATGCTTGGAGGGCCTCCACTTTGTGGCCAGC

- CTGTGGTCTGGCTCTGAGGCC TAGGCAACGAGCGACAGGGCTGCCAGTTGCCCTGGGTTCCCTTTGTGCTGCTGTGTGCCCTCC
TCTCCTGCCGCCCTTTGTCTCCGCTAAGAGACCTGCCCTACCTGGCCGCTGGGCCCGTGACTTTCCCTTCTGCCAGGAA
AGTGAGGCTCGGCTGGCCCCACCTTCCCTGTCTGATGCCGACAGCTTAGGGAAGGGCACTGAACTTGCATATGGGGCTTAGCC
TTCTAGTCACAGCCTCTATATTTGATGCTAGAAAAACATATTTTAAATGGAAGAAAAATAAAAGGCATTTCCCCCTTCATCC
5 CCCTACCTTAAACATATAATATTTTAAAGGTCAAAAAAGCAATCCAACCCACTGCAGAAGCTCTTTTGTGAGCACTTGGTGGCAT
CAGAGCAGGAGGAGCCCCAGAGCCACCTTGGTGTCCCCAGGCTACCTGCTCAGGAACCCCTTCTGTCTCTGAGAACTCAAC
AGAGGACATTGGCTCACGCACTGTGAGATTTTGTCTTATACCTGCAACTGGTGAATTTATTTTATAAAGTCATTTAAATATC
TATTTAAAGATAGGAAGCTGCTTATATATTTAATAATAAAGAAGTGACAAGCTGCCGTTGACGTAGCTCGAG
417
- 10 MSVPAFIDISEEDQAAELRAYLKSGBEISEENSEGGLHVDLAQII EACDVCLKEDDKDVESVNSVLSLLILEPKQAEALIE
SLCEKLVKPREGERPSRLQLLSNLFHGMKNTPVRYTVYCSLIEVVASCGAIQYIPTELDQVRKWI SDWNLTTEKKHTLLRL
YEALADCKKSDAASKVMVELLGSYTEDNASQARVDAHRCIVEPLKDPNAPLFDHLLTLKPVKFLEGLIHDLLTIFVSAKLASY
VKFYQNNKDFIDSLGLLHEQNMAKMRLLTFMGMAIENKEISFDTMQQELQIGADDVEAFVIDAVRTKMVYCKIDQTORQVVVSH
SHRTFGKQRWQQLYDTLNAWKQNLNKVNSLLSLSDT
418
- 15 418
GGGTCGGCGTGGTCTTGGGAGTGGAGTGCCGCTGTGCCGGGGCTGCACCATGAGCGTCCCGGCTTCATCGACATCAGTGAA
GAAGATCAGGCTGCTGAGCTTCGTGCTTATCTGAAATCTAAAGGAGCTGAGATTTCAGAAGAGAACTCGGAAGGTGGACTTCAT
GTTGATTTAGCTCAAATTATTGAAGCCTGTGATGTGTCTGAAGGAGGATGATAAAGATGTTGAAAGTGTGGTGAACAGTGTG
GTATCCCTACTCTTGATCTCGAACCAGACAAGCAAGAAGCTTTGATTGAAAGCCTATGTGAAAGCTGGTCAAATTTCCGCGAA
20 GGTGAACGCGCTCTCTGAGACTGCAGTTGTTAAGCAACCTTTTCCACGGGATGGATAAGAATACTCCTGTAGATACACAGTG
TATTGCAGCCTTATTGAAGTGGTAGCATCTGTGGGGCCATCCAGTACATCCCAACTGAGCTGGATCAAGTTAGAAAATGGATT
TCTGACTGGAATCTCACCCTGAAAAAAGCACACCCCTTTAAGACTACTTTATGAGGCACTTCGGGATTGTAAGAAGAGTGAT
GCTGCTTCAAAAGTCATGGTGAATTTGCTCGGAAGTTACACAGAGGACAATGCTTCCAGGCTCGAGTTGATGCCCACAGGTGT
ATTGTCGAACCATTTGAAAGATCCAAATGCATTTCTTTTGACCACCTTCTTACTTTAAACCAGTCAAGTTTTTGAAGGCGAG
25 CTTATTCAATGATCTTTTAAACCATTTTGTGAGTGCTAAATTTGGCATCATATGTCAGTTTATCAGAATAATAAAGACTTCATT
GATTCATTTGGCCTGTTACATGAACAGAATATGGCAAAATGAGACTACTTACTTTATGGGAATGGCAATAGAAAAATAAGGAA
ATTTCTTTTGACACAATGCAGCAAGAACTTCAGATTGGAGCTGATGATGTTGAAGCATTTGTTATTGACGCCGTGAAGAACTAAA
ATGGTCTACTGCAAAATTGATCAGACCCAGAGAAAAGTAGTTGTGTCAGTCATAGCACACATCGGACATTTGGAAAACAGCGGTGG
CAACAACCTGTATGACACACTTAATGCCTGGAAACAAAATCTGAACAAAGTGA AAAACAGCCTTTTGAGTCTTTCCGATACCTGA
30 GTTTTATGCTTATAATTTTGTCTTTGAAAAAAAGCCCTAAATCATAGTAAGACATTATAAACC AAAAAA
419
- MSANATLKPLCPILEQMSRLQSHSNTSIRYIDHAALLHLASLLGLVENGVLFFVVGCRMRTQTVVTTWVHLALS DLLLASASL
PFFTYFLAVGHSWELGTFCKLHSSIFFLNMFASGFLLSAISLDRCLQVVRPVWAQNHRTVAAAHKVCLVLWALAVLNTVPYFV
FRDTPISRLDGRIMCYYNVLLNPGPDRDATCNSRQAALAVSKFLLAFLVPLAIIASSHAAVSLRLQHRGRRRPRGRFVRLVAADV
35 AAFALCWGPYHVFSLLEARAHANPGLRPLVWRGLPFVTSLAFFNSVANPVLYVLTCPDMLRKLRRSLRTVLESVLVDDSELGGA
GSSRRRTSSTARASPLALCSRPEEPRGPRLGLWLLGSCAASPQTGPLNRALSSTSS
420
- CAGCCTCCCTCTCCACCTCTGTCTGCCCGCTGCCCTCTTGTCTAGCTGCTGTCAGGAGCTGACTGCCTCCAGGGCTGGAATCCT
GTGCTCCCTCTGTGCCAGAGCCCCAGATGTCGGCCACGCCACTGAAGCCACTCTGCCCCATCCTGGAGCAGATGAGCCG
40 TCTCCAGAGCCACAGCAACACCAGCATCCGCTACATCGAACCCGCGCGCTGCTGCTGCACGGCTGGCTCGCTGCTGGCCCT
GGTGGAGAAATGGAGTCATCTCTTCTGTGGTGGGCTGCCGATGCCCGACACCGTGGTCACCACTGGGTGCTGACCTGGCGCT
GTCCGACCTGTTGGCCTCTGCTTCCCTGCCCTTCTTCACCTACTTCTTGGCCGTTGGGCCACTCGTGGGAGCTGGGCACCACTT
CTGCAAACTGCACCTCCTCCATCTTCTTCTCAACATGTTGCCAGCGGCTTCTGCTCAGCGCCATCAGCTGGACCGCTGCCCT
GCAGGTGGTGGCGCGGTGTGGGCGCAGAACCCGACCGTGGCGCGCGCACAAAGTCTGCTGGTGTCTTGGGCACTAGC
45 GGTGCTCAACACGGTGGCCTATTTCTGTGTTCCGGGACACCATCTCGCGGCTGGACGGGCGCATTTATGTGCTACTACAATGTGCT
GCTCCTGAACCCGGGGCTGACCGGATGCCACGTGCAACTCGCGCCAGCGGGCCTGGCCGTGAGCAAGTTCTGCTGGCCTT
CCTGGTGGCGCTGGCGATCATCGCTCGAGCCACGCGCGCTGAGCCTGCGGTTGACGACCGCGCGCGCGCGCGCGGCGGCGG
CTTCGTGCGCTGGTGGCAGCGCTGCTGGCGCCTTCGCGCTCTGCTGGGGCCCTACCACGTGTTAGCCTGCTGGAGGCGCG

GGCGCAGCAAACCCGGGGCTGCGGCCGCTCGTGTGGCGGGCTGCCCTTCGTCACCAGCCTGGCCTTCTTCAACAGCGTGGC
CAACCCGGTGCTCTACGTGCTCACCTGCCCGACATGCTGCGCAAGCTGCGGCGCTCGCTGCGCAGCGTGCTGGAGAGCGTGCT
GGTGACGACAGCGAGCTGGGTGGCGGGGAAGCAGCCGCCCGCCGACCTCCTCCACCGCCCGCTCGGCCCTCCCTTTAGC
TCTCTGCAGCCGCCGGAGGAACCGGGGCCCCGCGCTCTCCTCGGCTGGCTGCTGGGCAGCTGCGCAGCGTCCCGCAGAC
5 GGGCCCCCTGAACCGGGCGCTGAGCAGCACCTCGAGTTAGAACCCGGCCACGTAGGGCGGCACTCACACGCGAAAGTATCACC
AGGGTGCCCGGTTCAATTGATATCCGGACTCTGCGCCAGTGATCAAAGTCCGAGGGGCGGACCCAGGCACCTGCATTTTA
AAGCGCCCCGGGAGACTCTGAATCTTTTTCAGAAACAGTGAGTTAAAGCAGTGCTTCTCAAACCTTGATGTGCTGTGAATCAC
CTAGGGGTCTTGTAAAGTGAGTCTGATCCAGGAGGCCGGGGCCGGTACTGAGAGTCTGCACCTAACAGCTCCAGGCCGAG
AAGCCAGTGCGGCAGGTTACAGGCGAGGCTGGAGTAACAAAGTGAAGTAACTCGTAATAGACTTCCCACTCTAGGGCAGTGGA
10 GTCGGAAGGGCACACGGGTGCGTCTCCCGGAGTTAGTTTACCAGATGATGGGGGAGGGGGAAGGAGTTTATGTTAAAC
CATCCATGTATTTTGGAGAAGAGAGAGGAAAGTTTGAGAAGCACTGTTCCAGCCTGCCCTCTTCACTTTAGCCAATGCTTACT
GCGCTAGACGCTTCATCCCAATCTTAAGGGGAGCTTCTATTTAGCCAGTCTTTACAGCTGAGCACATTTGCGCTCAGGGAGG
TTAAGTGACTTGCCAGTTTACAGGCTAACGACCACAGGCTCTGCACTCTAACCTAGGCATCATAGCTCAATGACTCTCTGG
TGAGCGAGGACATTTCTTGACCTACTCGAGGACTTAAGATGCTACCTTGTGACCCAGCACTGCCCAAAGTGCTTCCAAGGCAG
15 AAGCAGCAGGGGATGGCGTGGTCAAGCACTCGGGAACCTGGGGCTAATCAAATCCAATGGGGAAATGACTAAAGTCTTTCGG
TCGTTAGAAGTTGAATGGGCACAGCAACTAAGACTACAGCACAGCTCATTTCTTAGCTAAGCGGACCAAGCTCCCTGTCGGC
CTGGTGTCTTGTGGGATCCCTCTGGGCACGTGTAATCCCAAGATCTGTGCAGCCCCGCTCCAGGCCACATGGGGCTGGGCAGC
TACCATTTCCCTTTTGCAGGATGGGAGGGGTAACCTTGACCTCTGACCTATCACTTCCACTGCACCCCGTCTCATTTCTCCACCT
GCCGTGGACTTGGGGTCAGAGACTGCTGTGTTTGTAGCTCTGCAGCCAGGGACCGAAAGTTGGTGTCAATGAATTTGCTTGG
20 TGGATGAAATGTCAGTGGAAGAAGCAGATGAGAACTCTTGAGATCTTGGTCTGTGTTTTTCTGCCACCAAGGCCAGGGTC
ACTGAAGGCCTGGCCACAGCAGGTGCTGAGCAAAGGGAACAGTGAGGTGCCAGCTAGCTGCAGAGCCACCTGTGTTGACAC
CTCGCCCCTGCTCCCTCCCATCCCTTCCCTTTTACTCATAGCACTTCCCCCATGGACACGTGGTGCATTTTGTGTTTATT
ATGTTTTCTCTCCATCAGAAAGCTCCTCGAGGGCAGGGACTTTGGTCTATTGTCTGTATTGCGCGTGCTAGGATTGTG
CCTGTATGCAACAGGCACTCAATAAATATTTTGTGTAGACTGG

25 421

MASTSTTIRSHSSRRGFSASSARLPVSRSGFSSISVSRSRGSGGLGGACGGAGFGSRSLYGLGSSKRISIGGGSCAISGGYG
SRAGGSYFGGAGSGFGGAGIGFGLGGAGLAGGFGGPGFVPCPPGGIQEVTVNQSLTPLNLQIDPAIQVRABEREQIK
TLNNKFASFIDKVRFLQONKVLDTKWTLLQEQQTKTVRQNLPLFEQYINNLRRLDSIVGERGRDLSELNMQDLVEDLKNK
YEDEINKRTAAENEFVTLKKDVAAYMNKVELQAKADTLDEINFLRALYDAELSQMQTHISDTSVVLSDMNNRNLDLDSIIAE
30 VKAQYEEIAQSRABAESWYQTKYEELQVTAGRHGDDLNTKQETAEINRMIRLRSEIDHVKKQCANLQAAIADAEQRGEMAL
KDAKNKLEGLDALQAKQDLARLLKEYQELMNVKLALDVEIATYRKLLGECECLNNGEGVQVNI SVVQSTVSSGYGGASGVG
SGLGLGGSSSYSGSLGVGGGFSSSSGRATGGGLSSVGGSSSTIKYTTTSSSSRSYKH

422

ATGGCCAGCACATCCACCACCATCAGGAGCCACAGCAGCAGCCGCCGGGTTTCAGTGCCAGCTCAGCCAGGCTCCCTGGGGTC
35 AGCCGCTCTGGCTTCAGCAGCATCTCCGTGTCCCGCTCCAGGGGAGTGGTGGCCTGGGTGGCGCATGTGGAGGAGCTGGCTTT
GGCAGCCGAGTCTGTATGGCCTGGGGGCTCCAAGAGGATCTCCATTGGAGGGGCGAGCTTGCCATCAGTGGCGGCTATGGC
AGCAGAGCCGAGGAGCTATGGCTTTGGTGGCGCCGGAGTGGATTGTTGTTTCGGTGGTGGAGCCGCATTGGCTTTGGTCTG
GGTGGTGGAGCCGGCTTGTGCTGGTGGCTTTGGGGGCCCTGGCTTCCCTGTGTGCCCCCTGGAGGCATCCAAGAGGTCACTGTC
AACCAGAGTCTCCTGACTCCCTCAACCTGCAAATTGACCCCGCATCCAGCGGTGCGGGCCGAGGAGCGTGAGCAGATCAAG
40 ACCCTCAACAACAAGTTTGCCTCCTCATCGACAAGGTGCGGTTCTTAGAGCAGCAGAACAAAGGTCTTGACACCAAGTGGACC
CTGCTGCAGGAGCAGGGCACCAGACTGTGAGGCAGAACCTGGAGCCGTTGTTTCAGCAGTACATCAACAACCTCAGGAGGCAG
CTGGACAGCATCGTGGGGGAACGGGTCTGCTGGACTCGGAGCTGAGAAACATGCAGGACCTGGTGGAGGACCTCAAGAACAAA
TATGAGGATGAAATCAACAAGCGCACAGCAGCAGAGAATGAATTTGTGACTCTGAAGAAGGATGTGGATGCTGCCTACATGAAC
AAGTTGAAGTGAAGCCAAGGCAGACACTCTTACAGATGAGATCAACTTCTGAGAGCCTTGTATGATGCAGAGCTGTCCAG
45 ATGCAGACCCACATCTCAGACACATCCGTGGTGTATCCATGGACAACAACCGCAACCTGGACCTGGACAGCATCATCGCTGAG
GTCAAGGCCAATATGAGGAGATTGCTCAGAGGAGCAGGGCTGAGGCTGAGTCTTGGTACCAGACAAAGTACGAGGAGCTGCAG
GTCAAGCAGCAGGACATGGGGACGACCTGCGCAACACCAAGCAGGAGATTGCTGAGATCAACCGCATGATCCAGAGGCTGAGA
TCTGAGATCGACCACTCAAGAAGCAGTGTGCCAACCTACAGGCTGCCATTGCTGATGCTGAGCAGCGTGGGGAGATGGCCCTC
AAGGATGCTAAGAACAAGCTGGAAGGCTGGAGGATGCCCTGCAAGAAGCCAAGCAGGACCTGGCCCCGCTGTAAGGAGTAC
50 CAGGAGCTGATGAACGTCAGCTGGCCCTGGACGTGGAGATCGCCACCTACCGCAAGCTGCTGGAGGGCAGGAGTGACAGGCTG

AATGGCGAAGGCGTTGGACAAGTCAACATCTCTGTAGTGCAGTCCACCGTCTCCAGTGGCTATGGCGGTGCCAGCGGTGTCGGC
AGTGGCTTAGGCTTGGGTGGAGGAAGCAGCTACTCTATGGCAGTGGTCTTGGCGTTGGAGGCGGCTTTAGTTCAGCAGCGGC
AGAGCCACTGGGGTGGCTCAGCTCTGTTGGAGGCGCAGTTCACCATCAAGTACACCACCCTCTCTCCAGCAGGAAG
AGCTACAAGCACTGA

5 423

MNGPALQPSSPSSAPSASPAAPRGWSEFCELHAAAAARELARQYWLFAREHPQHAPLRAELVSLQFTDLFQRYFCREVRDGRA
PGRDYRDTGRGPPAKAEASPEPGPPAAPGLPKARSSEELAPRPPGPCSFQHFRRSLRHI FRRRSAGELPAAHTAAAPGTPE
AAETPARPGLAKKFLPWSLAREPPPEALKEAVLRYSLADEASMSD GARWQRRLALRRAPGPDGPDRLVLELFDPPKSSRPKLQA
ACSSIQEVWRCTRLMPDNLYTFVLVKVDRDITIFEVGD EQQLNSWMAELSECTGRGLESTEAEHIPSALPESTSSSPRGSTD
10 SLNQGASPGGLLDPACQKTDHFLSCYPWFHGPISR VKAAQLVQLQGPDAHGVLVVRQSETRRGEYVLTFFNQGI AKHLRLSLTE
RGQCRVQHLHFP SVVMDLHHFQ RSPIPLECGAACDVLSSYVVVVSQPPGSCNTVLPFSLPHWDESLPHWGSELGLPHLSSS
GCPRGLSPEGLPGRSSPPEQIFHLVPSPEELANSLQHLEHEPVNRARDSDYEMDSSSRSHLRADINQYTPL

424

CCCCGGCCACCGCTCCGCCCGGCTGCCCGCCGGACTGTGCGGCCCGCGGTGGCGACGGCGGCCGCTGCAAAGTTTCCCCGG
15 CGCGCGCGCGCCGGGGCGCATCTCCCGCACTGTCAAGCGCTGGCGCGGAAATGATGAGGCGCTGGCCATTTTCCGAGCCC
GGGTTTCTCGCTGAGCCCCGCTCGAGCGAGCCGCGAGCGAGGACCGCGGGCGGAGAGGACCGCCAGGCGGGGGCCCG
CCCCCCCCCTCGGATTTCGAGGGCCCCGGGGCGCGCAGCCATGGGCGGGCGGCCAGAGCTCTGTCTCTCAGCCCGGC
CGCACCACTGGGTCTCCGCCATGAACGGGCTGCCCTGCAGCCCTCTCGCCCTCTTCCGCGCCCTCAGCCTCCCCGGCGGCG
GCCCCGGGGCTGGAGCGAGTTCTGTGAGTTGCACGCCGTAGCGCGGGCGGGAGCTGGCCCGCCAGTACTGGCTGTTCGCC
20 CGGGAGCATCCGACGACGCGCCGCTGCGCGCCGAGCTGGTGTGCTGCTGAGTTACCGACCTCTTCCAGCGCTACTTTCGCCG
GAGGTGCGCGACGACGGCGCCGGCGCGACTACCGGGACACAGGCCGTGGGCCCCAGCCAAGGCCGAGGCGTCCCGGGAG
CCAGGCCCGGGCCCCGCCGCCCTGGCTGCCCAAGGCCCGAGCTCTGAGGAGCTGGCCCCCGCGCGGCCCGGGCCCTGC
TCCTTCCAGCACTTTTCGCCGAGCTCCGCCACATCTTCCGCCCGCGCTCGGCCGGGAGCTGCCAGCGGCCACACCGCTGCC
GCCCCGGGACCCCCGAGAGGCTGTGTAGACCCCCGCCCGGCTGGCTGGCCAAGAAGTTCTTCCCTGGAGCTGGCCCCG
25 GAGCGCCACCCGAGGCGCTGAAGGAGGCGGTGCTGCGCTACAGCTGGCCGACGAGGCTCCATGGACAGCGGGGACGCTGG
CAGCGCGGGAGGCTGGCGCTGCGCCGGGGCCCCGGGCCCGATGGCCCCGACCGCGTGTGGAGCTCTTCGACCCACCCAAGAGT
TCAAGGCTCAAGCTACAAGCAGCTTGCTCCAGCATCCAGGAGTCCGGTGGTGACACGCTTGAGATGCCTGACAACCTTTAC
ACCTTTGTGCTGAAGGTGAAGGACCGGACAGACATCATCTTTGAGGTGGGAGACGAGCAGCTGAATTCATGGATGGCTGAG
CTCTCGGAGTGCACAGGCCGAGGGCTGGAGAGCACAGAAGCAGAGATGCATATTCCTCAGCCCTAGAGCTAGCAGTCCAGC
30 TCCCAAGGGGACGACAGATTCCCTTAACCAAGGTGCTTCTCTGGGGGCTGCTGGACCCGGCTGCCAGAAGACGGACCAT
TTCTGTCTGCTACCCCTGGTTCCAGGCCCATCTCCAGAGTGAAGCAGCTCAGCTGGTTACGCTGCAGGGCCCTGATGCT
CATGGAGTGTCTGTTGCGGAGAGCGAGACGCGCGGTGGGGAATACGTGCTCACTTTCAACTTTCAGGGATAGCCAAGCAC
CTGCGCTGTGCTGACAGAGCGGGGCGAGTGGCGTGTGACGACCTCCACTTTCCCTCGGTGCTGGACATGCTCCACCACTTC
CAGCGCTCGCCATCCACTCGAGTGGGCGCGGCTGTGATGTCCGGCTCTCAGCTACGTGGTAGTCTCTCCCAACCACCA
35 GGTCTCTGCAACACGGTCTCTTCCCTTTCTCCCTTCTCACTGGGATTCAGAGTCCCTTCTCACTGGGGTTCAGAGTTGGGC
CTTCCCCACCTTAGTTCTTCTGGCTGTCCCCGGGGGCTCAGCCAGAGGGTCTCCAGGGCGATCTCACCCCCGAGCAGATC
TTCCACCTGGTGCCTTCGCCGGAAGAACTGGCCAACAGCTGCAGCACCTGGAGCATGAGCTGTGAATCGAGCCCGGGACTCG
GACTACGAAATGGACTCATCTCCCGAGCCACCTGCGGGCCATAGACAATCAGTACACACCTCTCTGACCAGTGAAGAAATTC
AGGCTCAACAGCTGCCCTTGAGGAGCACAGGCAAGTGTGAACCTGTGAATGTAATTGATCTTCTTCTTCCAGAGAAAG
40 ATTTAAGGGACACTGTTAACTGCTCGTGCCAGTTTGAAGTGACCTTCTATTAGCGCTGTTGAAGGGCCCTCTGTAGGTTTC
ATCTATCCACCTGGCTTCTCCTTATGTTTACAGATGTAGTTCTTGTAGAGGATGCCGCTAGCTCCTGCCGGGGTCCCTAT
GCCAGTCCCCGTACTCTTAGAGAAAGGAGTTGGGGTGAAGGCCAGAGCTGGCAGTGGAACTTGTCTCTTTTCACTGACA
CTGTACAGCGGATGACAGACTTTCTACGGGAGGAGGGGGGATCATCAGGAAGCCAGAACACTAACAGCGGTTCTCCCAT
CTACCGTCACTCCATGGCAGGTCTGCTGTGTCACACCACAGATGACCACATCTAATCTGCTTCTACTCTCAGCTTTAGGA
45 CAAAAGCTCTGTGAGGCGACAAGCTGAAGGTCAAAAACATTTTAAACATTTTACCTCAGACTAATTTCTTTAAAGGATTCAG
GTTCAAACTTAACCACTGCTTATTTCACTGCTGCTTTCAACTAACACCATGCTATTTTGTAGTCAAGAACAGCTATGCAA
ACCTTACCTAATTTACAGTCTGAGCCAGCATGCTGGCTGTCTACTGCATCTCGGACAGTACCTGCCACTGAGTGGCCACT
GTCTTCTCTAAATGTCAAGAAGTGAAGTATGTACCCCTTTCAGGGAATTCAGGCAATTACTGAAATAGGAAGGTGGCAAGAAC
AGTTCTATCTGTTGCCCTTACGAATAAAAACTGGATTCTGGTTTACAGCAGCTTTACAGTGATAGTTAAATTAAGTGGGGCTA
50 GGGGAAAAGCAACCAAAAAGGAAAAAGGACTCCTAGGCCCTTTCTATTAAATCCTTCAGCAACAAGGCTGGCTTGGTGCCCTC

CAAGCATCTAATGGCTTATTAAATTATCCACAATTGGGTTTTAGGCTCCTTTTTTGACCCAAAATGGAAGCTGGGAATCTGGT
GCCATAACTAATGAGAACTCCTTTAATACCCCAACAATCAGTGTCTGTCTTACTGGCTACTGCTTCACTGGATTGAAAATCT
ATCTATCTCCTTGACACATGGGCACACACAATCTCCACCATCCAGGGAGGTCTGAATTCAAATCTCTATCTATCCAAGTGAT
ACAATTCATAGGGGGCTGGCTCCTCCAGAACCTGTCTGGAGGCTCAGAAAACGGGGGAGTGACAGTGGAGTCAGCTGCCTTG
5 GGTGCCAGCAGAGCCATTCAGTACAACCCCCAGGCTCACAGCAGTGGCTTCTAGGAACTGGGAGTTTATAGTCAGCTTTACAGA
TACATCGATCAGAGGCTAAATGAAACCTCAGCCTAAAACCTCATAGGACTGACTGCCTGGGAGGAGGGTTAGGTCTGCCTTCTC
CACTTATACTTAGTCTCTGTGCTCCAAGAGGTCAAATTTTTGCTTCTAGAAATTTCTTGGGGTCTTTACAGGGTGGGGGAACA
AACCCTATGCACTTTTCTTTTTTTTTTTGAGATGGAGTTTCTTCTGTCAACCGGGCTGGAGTGCAGTGGTGAATCTTGGCT
CACTGCAACCTCCACCTTCTTGGTTCAAGCGATTCTGCCTCGACCTCTCAAGTAGCTGGGATTACAAGCACCAGCCACCATGCC
10 TGGCTAATTTGTATTTTGTAGTAGAGACAGGTTTACCATGTTGGCCAGGCTGGTCTCGAATGTCTGACCTCAGGTGATCCAC
CCGCTTTGGCTCTCCAAAGTGTGGGATTACAGGCGCGAGCCACCGCGCCAGCCTACACCCTTTTAGTACCAACACTCTTGG
GTGATTTTCATGGACCTTAAAGCAGACCTGACACTGATCAGATTGTCAGTCCATTTTAAAGACACCTGTCTTTATTTCTCTCA
AGTCAAGCAGCTTTCTCTGGAAAATGAATGCTAATTAGTGTGAACCAAAAGAGTAAGTAAGAGTCTGAAGTTTTTTTAAAGGAG
AAAGCTTATTATGGAAGTCACTGGTCTCCCTCCGACAGGAAAGGTACCCAGTAGATAATGAACCAAAATTAAGTTCCCTCC
15 CTCCAGCCAGAAGTTAAACATCTGGGATATGACGTCTTCATGCCAGGGGCACTCATTTCTTAGCAGCCTCTCTACATACATCTC
TCAGGTGGTGCCAGAGGCACACCAGGTAGAGCAAACTTAGCAGCTCTGACTAACAGGCTGCAAGTGCAAGTTTCAGATTCTGT
GGCAGAGATTTGGAGGGCACCCACGTCCAGACTGCTTCCCGTCAAGTTACCAGGACAGCTCAAAAACATGCTGACAGAAAAC
CCCATGGCTCTAGGAAAAGTGACACTAAGCCAACACCTTTCTTTATGTGGGAGCAAAATCAGCTGATGAAGGGGTGGGCACCA
TTGTGGGGCAGGCACCCCACTGGCTGCAGCTAGCCACCATAGGCACAGCAGCATCCCACTCTCCTTCCAGTCTTGACCAGG
20 CCCCAGCCGGCAACTTCTACCGAGAGCCATGGCTCAACACCAAACTGGACAGTAGACATCATGATCCCTCCAGTTAGCTCTAAT
TACAGACCCCACTAGTACAGCTTGACAGCTCCCGGCACCATCCCTTCTCATCTGACTTATTGAATTTTACAACTAACAGT
CACCAGCACCAGAAATTAAGTCAACTAACCTGCCTTGAATTTTAAACAGCAATCCATATGGCTTTATCTGGTATAAATCTTC
TGCTTTTGATCATTTCTGACCGTAGGAAAAAGGAATAGCAATCTTAAATCTTGGGCGAGAGAACATATTTTACATAACA
GTTTCTTAACTAAAGTCAAGGCCTTGGACTCTTCCCTGAGGGTTGCCTGAAATTCCTTCATGCTTTCTATTACAGACTAATTC
25 CTTTACTGCAAAATGTGTTAGCTCTAACATCTCCACAAGCTAAAGGAACCTGCAAGTATATTAACAAGGACACATCTGACATCC
TGTGTTTGGTTAAATATACAGCAGATTGTGATAACATAAAGTGGATCCATCTGTATCATTTATAGGCAAAAGGTATTTGGCAA
ATTTTTATGTATGGTTTTATGTACTGTACAAGTAACTTATTCTTGAATAATGCAAAATTTGCTATAATGTACAAATGCTATAT
GTGAATTAAGTTTCCAAAATCTTGAAAAAATTT
425

30 MAAAPIQONGHTGVPIIDLDPDSRKRPLEAPPEAGSTKRTNTGEDGQYFLKVLIPSYAAGSIIGKGGQITIVQLQKETGATIK
LSKLSKSKDFYPGTTTERTVLIQGTVEALNAVHGFIAEKIREMPQNVAKTEPVSILOPQTTPNPDRIKQTLPSPTTTKSSPSDP
MTTSRANQVKIIVPNSTAGLIIGKGGATVKAVMEQSGAWQLSQKPDGINLQERVTVSSEPEQNRKAVELIIQKIQEDPQSGS
CLNISYANVTGPVANSNPTGSPYANTAELVPTAAAAAGLGHANLAGVAAPPAVLGFTGNDLVAITSALNTLASYGYNLNTLG
LGLSQAAATGALAAAAANPAAAAANLLATYASEASASGSTAGGTAGTFALGSLAAATAATNGYFGAASPLAASAILGTEKST
35 DSKDVVEIAVPENLVGAILGKGGKTLVEYQELTGARIQISKKGEFVPGTRNRKVTITGTPAATQAAQYLITQIRITYEQGVRAA
NPQKVG
426

40 GAATTCGACAAAAACAAAGGGAGAACCTTCTCCCGGTAGCAGCGGCAGGAACTGCAACATGATGGCGGCAGCTCCCATCCAG
CAGAACGGGACCCACACTGGGGTTCCTATAGACCTGGACCCGCCGACTCGCGGAAAAGGCCGCTGGAAGCCCCCTGAGGCC
GGCAGCACCAGAGGACCAATACGGGCGAAGACGGCCAGTATTTTCTAAAGGTTCTCATACCTAGTTATGCTGCTGGATCTATA
ATTGGGAAGGGAGGACAGACAATTGTTTCAAGTTCGAAAAAGAACTGGAGCCACCATCAAGCTGTCTAAGCTGTCTAAGTCCAAA
GATTTTACCCAGGTACTACTGAGCGAGTGTGCTTGATCCAGGGAACGGTTGAAGCACTGAATGCAGTTTCATGGATTATTGCA
GAAAAAATTCGAGAAATGCCCCAAATGTGGCCAAGACAGAACCACTGAGCATTTCTACAACCCAGACCCAGTAAATCCAGAT
CGCATCAACAAACATTGGCATCTTCCCCAAGTACCACCAAGTCTCTCCATCTGATCCCATGACCACCTCCAGAGCTAATCAG
45 GTAAAGATTATAGTTCCCAACAGCAGCAGGCTGATAATAGGGAAGGAGGTGCTACTGTGAAGGCTGTAATGGAGCAGTCA
GGGGCTTGGGTGCAGCTTTCCAGAAACCTGATGGGATCAACTTGCAAGAGAGGGTTGTCACTGTGAGTGGAGAACCTGAACAA
AACCAGAAAGCTGTTGAACCTATCATCCAGAAGATACAAGAGGATCCACAAAGTGGCAGCTGTCTCAATATCAGTTATGCCAAT
GTGACAGGTCCAGTGGCAATTCATCCAAACCGGATCTCTTATGCAAACTGCTGAAGTGTACCAACTGCTGCAGCAGCT
GCAGGGCTATTAGGACATGCTAACCTTGTGGCGTTGCAGCCTTTCCAGCAGTTTATCTGGCTTCACAGGCAATGACCTGGT
50 GCCATCACCTCTGCACTTAATACATTAGCCAGCTATGGATATAATCTCAACACTTTAGGTTTAGGTCTCAGTCAAGCAGCAGCA

ACAGGGGCTTTGGCTGCAGCAGCTGCCAGTGCCAACCCAGCAGCAGCAGCAGCCAAATTTATTTGGCCACCTATGCCAGTGAAGCC
TCAGCCAGTGGCAGCACAGCTGGTGGTACGGCGGGGACATTTGCATTAGGTAGCCTGGCTGCTGCTACTGCTGCAACCAATGGA
TATTTTGGAGCTGCTTCTCCCTAGCTGCCAGTGCCATTTCTAGGAACAGAAAAGTCCACAGATGGATCCAAAGGATGTAGTTGAA
ATAGCAGTGCCAGAAAACCTTAGTTGGTGCATATCTTGGCAAGGAGGGAAAAACATTAGTGGAATACCAGGAGTTGACTGGTGCA
5 AGGATACAGATCTCCAAAAAGGAGAAATTCGTACCTGGCACAAGGAATCGGAAGGTAACCATTTACTGGAACACCAGCTGCAACA
CAGGCTGCTCAATATTTAATTACACAAAGGATCACATATGAGCAAGGAGTTCCGGCTGCCAATCCTCAGAAAAGTGGGTTGAGTG
CCCCAGTTACACATCAGATTTGTTTAAACCCCTCCTTTACCCCATTTTCAAGAAGGATGTACTGTACTTTGCAGAAAGTGAAGTTT
TTCTGTTATTAATATATAATTTATGCAATGAATGCGACTATGTTGACAATGTGTATATGTAAATAATATGTGTTTTACCAGATG
TTTCATAGAAAGAAATTTTTCTTGATCTGTTTTGTTCTCTATACTTTGTGTTGTATATTTGTGAGAGGTTTCTAGTGTAAG
10 ATTTAAGCCTGCCATTTTACCAGCATTATGTAGTTTAAATGATTGAATGTAGACAGGGATATGCGTATAGTTTTTCAGTATTAGT
TGTAGATAACACTAAATTAACACTAGTTAGTTAGTTAGTATGTTGGGTCAGTGACCTAAAATGGAGTGAGGCCAAAGCACTGTCC
TGTAAGTCTTACTCTCTGCTTAGGGCAGCTGAAGTAGGAAACAATTTTGAATAAAGTTTAAATTTAAATATGATCAAAAA
GCAATATAGTTGCAATAAAGCACTGTAAATATTTAAAGGTTAAACACTGTGGAATAATATTGGTAACTTTACAGATCAATA
AAAGCACCTGTTCTCCATCTGAACTAGACAATGGAATAATGCTGCATGCTGCCATGGCCCATTTCTCATCATTTGTGAAGTTCA
15 ACAAAGTTCTCACATGGAGTCCCACCTCTTCAGAGGTTGCACATTTGTTTTTAAGACTGAATTCACTACTGATCCCATCGCCT
GGCCGAGACAGTCATTACTCCATTAACATCCTCACTGTTTACAGACATAACTGTGGTACAGGATTGGAATTTATAAAACAAAGT
GAAAGTGCCAACAAATTTATGATAGCTGATAATGTTTCATATCTGCAACTGCTTGATAAGTATGTTGCATTTTAAGAGCTTATA
ATTGTGTATAATTTGTTAACTAGAACTTATAGTATTTGTAATGTAGATTTTACTGTGAAGCTATCTGTGATTTAGCTGTT
TGCTCCCATGATGGAGTCTTTGCAGCATGGCGCTAGCAGCCAATGCAGTTTCTAATACTCGGTAATTTGCATGTTTTGTGGAGC
20 ATTTTATGTACCAACCAGACAGTATTTCTGTCATGCTTATTTAGAAGAGGCAGCTTATCTTGAGAGGTAGTGTATCTACCT
TTGTGAGGCTTTTTGACAGGTCATTTGAGAGTAAGCCTTTGTTCCCAAGACCAACAACCTGTCACCTCTTCTGTACCTCTCCT
GAGTGCCAACCTGTCCAGGCCATTTGACACACCATCTGTTAACCTCTGAGTTTGGCCACTCAAGGCCACTCATAGGGGCATCCTA
GCCCTGTGCACTCAGCACTCATAGGATCATCCAGACTCTCATGCGGCATGCAGTCTAATCATGACAAATATGCTGCTACTCTG
ATATCTGGCTGAGCAACTGAATTACAAAAGAGAATTACTTCCATCTCAACTTCAACCCATTGATTACGTCCATCCTAGCAAGCT
25 AAATGGCATCCCAGCTGCTCCTTTCTGTGCAACCAATTAAAGAACAATGAGTGTGATGCTCCATGTCTGAATTTCTGTCAGCCT
CTCTCTGAAGTGTGATCTTTGTCTCATGAACCTTTCCCTTTTGTTCATTGAACATATGAGACTCTTCATTTTCATATTGATTAC
TGTGCAATTTACTTTTGGACATTGAGAACTTGAAATTTTCTGTATCCCTTCCCTTCCACTATTAATAATTCATTTCTGTCA
AACTGTAAGAGTAGACTCATTTTCTTTTCTTTTAGTTTAAACATTGGAGCTGTTATTTTCATTTAGAGTTCTCTATCTCTAAATA
TTTATTTAGAGAATGATTTTAAAGGGAATGATATGCTTGTTTAAATGAAAGAGAAAAGCTGTAGTAACTGTGTTAATTGGTA
30 ATGACTATTTATCGTCGATACCTGTAGCTGTGTAAGTTTGCACAAATAGTGTATCTCGTGAATCAGTGCTTAGCATTTGCCGC
TATTATATTTACTCATTTTATCATTTATAAATGTGTTTAGTTTCATCATGTAGCATCAAAA

427

MAFIRKKQQEQQLQLYSKERPSLLLLNLBYYFEQHRANHLHKGSHHERKIRGSLKICSKSVIFEPDSISQPIIKIPLRDCIK
IGKHGENGANRHFTKAKSGGISLIPSQVYFIKEHNVPYKIERGKMEYVFELDVPKVEDVVELLQLHRASCLDKLGDQTAM
35 ITAILQSRILARTSFDKNRFQNISEKLHMECKAEMVPLVTNPGHVCITDTNLYFQPLNGYKPKPVVQITLQDVRRRIYKRRHGLMP
LGLEVFCTEDDLCSDIYLFYEPQDRDDLIFYIATYLEHHVAEHTAESYMLQWQRGHLNSNYQLHLNLADRSCNDLSQYPVF
PWI IHDYSSSELDLSNPGTFRDLSKPVGALNKERLERLLTRYQEMPEPKFMYGSHYSSPGYVLFYLVRIAPEYMLCLQNGRFDN
ADRMFNSIAETWKNCLDGATDFKELIPEFYGDVDFLVNSLKLDLGKRQGGQMVDDVELPPWASSPEDFLQKSKDALESNYVSE
HLHEWIDLIFGYKQKGSDAVGAHNVFHLPLTYEGGVDLNSIQDPDEKVAAMLTOILEFGQTPKQLFVTPHPRRITPKFKLSQTS
40 YNASMADSPGEESFEDLTEESKTLAWNITKLQLHEHYKIHKEAVTGITVSRNGSSVFTTSQDSTLKMFSKESKMLQRSISFSN
MALSSCLLLPGDATVITSSWDNNVYFYSIAGRRQDTLMGHDDAVSKICWHDNRLYSASWDSTVKVWSGVPAEMPGRKHHFDL
LAELEHDVSDVTISLNAASTLLVSGTKEGTVNIWDLTTATLMHQIPCHSGIVCDTAFSPDSRHVLTSTGTDGCLNVIDVQTMGLI
SSMTSDEPQTCFVWDGNSVLSGSQSGELLVWDLGAKISERIQGHTGAVTCIWMNEQCSSIITGGEDRQIIIFWKLY

428

GAATTCGCCTCGATGGCGTTTATCCGGAAGAAGCAGCAGGAGCAGCAGCTGCAGCTCTACTCCAAGGAGAGATTTCTCTGCTG
CTGCTTAACCTGGAGGAGTACTACTTTGAACAGCATAGAGCCAATCACATTTTGCACAAGGGCAGTCACCATGAAGGAAAAATC
AGAGGCTCCTTAAAAATATGTTCAAAATCGGTGATTTTGAACAGATTCAATATCCAGCCCATCATCAAGATTCTTTGAGA
GACTGTATAAAAAATAGGAAAGCATGGAGAAAATGGAGCCAAATAGACACTTCACAAAGGCAAAATCTGGGGGTATTTCACTCATT
TTCAGTCAGGTATATTTTCAATTAAGAACATAATGTTGTTGCACCATATAAAATAGAAAGGGGCAAAATGGAATATGTTTTGAA
50 TTGGATGTTCCCGGAAAGTGAAGATGTTGTGGAGACGTTGCTTCAGCTTCACAGAGCATCCTGCCTTGACAAATGGGTGAC

CAAACCGCCATGATAACAGCTATTTTGCAGTCTCGTTTAGCTAGAACATCATTTGACAAAAACAGGTTCCAAAACATTTCTGAA
AAGTGCACATGGAATGCAAAGCAGAAATGGTGACGCCCTCTGGTGAATAATCCTGGACACGTGTGCATCACGGACACAAACCTG
TATTTTCAGCCCCCAACGGCTACCCGAAACCTGTGGTCCAGATAACACTCCAAGATGTCCGCCGATCTACAAAAGGAGGCAC
GGCCTCATGCCTCTGGGCTTGGAAAGTATTTTGCACAGAAGATGATCTGTGTTCCGACATCTACCTAAAGTTCTATGAACCTCAA
5 GATAGAGATGATCTCTATTTTACATTGCCACATACCTAGAGCACCATGTGGCGGAGCACACTGCTGAGAGCTACATGCTGCAG
TGGCAGCGTGGACACCTTTCCAATATCAGTACCTCCTTCACTCAACAACCTGGCCGACCCGAGCTGCAACGACCTCTCCAG
TACCCTGTGTTTCCATGGATAATACATGATTATTCAGCTCAGAACTAGATTGTCAAATCCAGGAACCTTCCGGGATCTCAGT
AAGCCAGTAGGGGCCCTAAATAAGGAACGGCTGGAGAGACTACTGACACGCTACCAGGAAATGCCTGAACCAAAAGTTTACGTAT
GGGAGTCACTACTCTTCCCGGGTTATGTACTTTTATCTTGTAGGATTGCACCAGAGTATATGCTGTGCCATGCAAGTGA
10 AGATTTGATAATGCAGATAGAATGTTCAACAGTATTCAGAACTTGGAAAACTGTCTGGATGGTGAACGGGATTTTAAAGAG
TTAATTCAGAAATCTATGGTGTATGTGTAGCTTTCTAGTCAATAGCCTGAAGTTGGATTGGGAAAGAGACAAGGAGGACAG
ATGGTTGACGACGTGGAGCTTCCCTTGGGCTTCCAGTCCCGAGGACTTTCTCCAGAAGAGCAAAGATGCATTGGAAAGCAAT
TATGTGTCTGAACACCTTCACGAGTGGATTGATCTAATAATTTGGCTACAAACAAAAAGGGAGTGATGCAGTTGGGGCCCATAT
GTATTTCAATCCCTGACCTATGAAGGAGGTGTAGACTTGAACAGCATCCAGGATCCTGATGAGAAGTAGCCATGCTTACGCAA
15 ATCTTGAATTTGGGCAGACACCAAAACAACTATTTGTGACACCACATCCTCGAAGGATCACCCCAAAGTTTAAAGTTTGTCC
CAGACCTCCAGTTATAATGCTTCTATGGCAGATTCCCCAGGTGAAGAGTCTTTTGAAGACCTGACCGAAGAAACAAAACTG
GCCTGGAATAACATCACCAAACTGCAGTTACACGAGCATAAAAAACCAAGAGCAGTTACTGGAATCACGGTCTCTCGC
AATGGATCTTCAGTATTCACAACATCCCAAGATTCCACCTTGAAGATGTTTCTAAAGAATCAAAAATGCTACAAAGAAGTATA
TCATTTTCAAATATGGCTTTATCGTCTTGTTTTACCTTTACCAGGAGATGCCACTGTCATAACTTCTTCATGGGATAATAATGTC
20 TATTTTATTCATAGCATTTGGAAGACGCCAGGACAGCTTAATGGGACATGATGATGCTGTAGTAGAATCTGTTGGCATGAC
AACAGGCTATATTCGATCGTGGGACTCTACAGTGAAGGTGTGGTCTGGTCTCCTGCAGAGATGCCAGGCACCAAAAGACAC
CACTTTGACTTGTGCGCGAGCTGGAACATGATGTCAGTGTAGATACAATCAGTTTAAATGCTGCAAGCACACTGTTAGTTTCC
GGCACCAGAAAGGACAGTGAATATTTGGGACCTCACACGGCCACCTTAATGCACCAGATTCCATGCCATTTCAGGGATTGTA
TGTGACACTGCTTTTAGCCAGATAGTCGCCATGTCTCAGCACAGGAACAGATGGCTGTCTTAATGTCTATGATGTGCAGACA
25 GGAATGCTCATCTCCTCCATGACATCAGATGAGCCCCAGACGTCTTTGTCTGGGATGGAAATTCGGTTTATCTGGCAGTCAG
TCTGGTGAATGCTCGTTTGGGACCTCCTTGGAGCAAAATCAGTGAGAGAATACAGGGCCACACAGGTGCTGTGACATGTATA
TGGATGAATGAACAGTGTAGCAGTATCATCACAGGAGGGGAAGACAGACAAATATATTTCTGGAAATTCAGTATTAAGTGCCT
TTTCTCTCTCTGAATATTAATTTGAATCTATTTAATGCATTTTAAACCAAACCTTTTAAACCGGACTGGTGAATGTGCAATGTT
AGTAATTAGAAGTTTACCACATGGAAAAATTTGTGGTTTTAAACTTTCTAAATCATGGTGACTTCATTGAAAGCCATTAGTTGC
30 CATCTCTTATGGGCAGATAAAATGCGGCTGTGTTAGGAAAAACATGTTACACTGTAAGGCAGATGATCGTCCCGTATGATGAT
TGTCAGAAAGACAGGACTAAGTAGCAGAGAAATAGCTAAGAGATAAATTTGGGCTGGGGAACCTTGTGAGAAGCACTGAACCAATTA
AGAAATTTTCCAAGAAATGTGCAGTATCTCTGCTACTTCTGAATCTGTTTGTCTTCTTAATCTATCACAATTTGCCACCCAT
CGGGTTTTGGGTGTGTGTTTTCATAGCGTGGTTACTTTCTATAATGCTGTACCCAGATTCTAAGAACCTGGAGAGGATTAGCA
GTTCTTAGTAAGTTTACTGTGTATAGGAACGGTTTGTATTTTATTACAGCTATTTCATCTTTTCTACATTAAAAATATTTTCTC
35 TAAAAA

429

MLALRCGSRLGLLSVPRSVPLRLPAARACSKSGDPSSSSSSGNPLVYLDVDANGKPLGRVVLKADVVPKTAENFRALCTG
EKGFYKGSFTFHRVIPSFMCQAGDFTNHNGTGGKSIYGSRFPDENFTLKHVGPVLSMANAGPNTNGSQFFICTIKTDWLDGKH
VVFHGVKEGMDVVKKIESFGSKSGRTSKKIVITDCGQLS

40 430

GAATTCGGAGTTCCGGGCGCGCGACGTCAGTTTGAATTTCTGTGTTCTCCCCGCCGCTGTCGCCGCCGACCCGCGCCCGGA
TGCTGGCGCTGCGCTGCGGCTCCCGCTGGCTCGGCTGCTCTCCGTCCCGCGCTCCGTGCCGCTGCGGCTCCCGCGCCCGCG
CCTGCAGCAAGGGCTCCGGCGACCCGCTCTCTCTCTCTCCCGGAACCCGCTCGTGTACCTGGACGTGGACGCCAACGGGA
AGCCGCTCGGCCGCGTGGTGTGAGCTGAAGGCAGATGTCGTCCTCCAAAGACAGCTGAGAACTTCAGAGCCCTGTGCACTGGT
45 AGAAGGGCTTCGGCTACAAAGGCTCCACCTTCCACAGGGTGATCGTCTCCTTCACTGTGCCAGGCGGCGACTTACCAACACCA
ATGGCACAGGCGGGAAGTCCATCTACGGAAGCCGCTTCTGACGAGAATTTTCACTGAAGCACGTGGGGCCAGGTGTCTGT
CCATGGCTAATGCTGGTCTAACACCAACGGCTCCAGTTCTTCATCTGCACCATAAAGACAGACTGGTTGGAATGGCAAGCATG
TTGTGTTGGTTCACCTCAAAGAGGGCATGGACGTCGTGAAGAAATAGAATCTTTCCGGCTCTAAGAGTGGGAGGACATCCAAGA
AGATTGTCTACACAGACTGTGGCCAGTTGAGCTAATCTGTGGCCAGGGTGTGGCATGGTGGCAGCTGCAATGTCCATGCACC

411

CAGGTGGCCCGCTTGGGCTGTCAGCCAAGGTGCCGTGAAACGATACGTGTGCCACTCCACTGTCACAGTGTGCCTGAGGAAGGC
TGCTAGGGATGTTAGACGGAATTCC

431

MEPKRIREGYLVKKGSVFNTWKPMWVLLLEDGIEFYKKKSDNSPKGMIPLKGSTLTSPCQDFGKRMFVFKITTTKQQDHFFQAA
5 FLEERDAWVRDINKAIKIEGGQKFPARKSTRRSIRLPETIDLGALYLSMKDTEKGIKELNLEKDKKIFNHCFTGNVIDWLVS
QSVNRNQEGLMIASSLLNEGYLQPDGMSKSAVDGTAENPFLDNPDIFYFPDSGFFCEENSSDDVILKEEFRGVIIKQGCLL
KQGHRRKNWKVRKFIREDPAYLHYYPAGAEDPLGAHLRGCVVTSVESNSNGRKSEENLFEIITADEVHYFLQAATPKERT
EWIKAIQMASRTGK

432

10 GGCCACAGTCTGAGAGGAGTTGCCCTGAGAGTGACCTTTGCATCTGCCCTGTCCAGCCAGCATGGAACCAAAGCGGATCAGAGAG
GGCTACCTTGTGAAGAAGGGGAGCGTGTTCATACGTGGAACCCATGTGGGTGTATTGTTAGAGATGGAATTGAATTCTAT
AAGAAGAAAAGTGACAACAGCCCCAAAGGAATGATCCCGCTGAAAGGGAGCACTCTGACTAGCCCTTGTCAAGACTTTGGCAAA
AGGATGTTTGTGTTTAAAGTACTACGACCAACAGCAGGACCACTTCTTCCAGGCAGCCCTTCTGGAGGAGAGAGATGCCTGG
GTTCCGGGATATCAATAAGGCCATTAAATGCATTGAAGGAGGCCAGAAATTTGCCAGGAAATCTACCAGGAGGTCCATTGCACTG
15 CCAGAAACCATTGACTTAGGTGCCTTATATTTGTCCATGAAAGACACTGAAAAAGGAATAAAAGAACTGAATCTAGAGAAGGAC
AAGAAGATTTTAACTACTGCTTCACAGGTAAGTGCCTGATTGATGGCTGGTATCCAACAGTCTGTTAGGAATCGCCAGGAA
GGCCTCATGATTGCTTCATCGCTGCTCAATGAGGGGTATCTGCAGCCTGCTGGAGACATGTCCAAGAGTGCAGTGGATGGAAGT
GCTGAAAACCCCTTTCTGGACAACCCGTGATGCCCTTCTACTACTTTCCAGACAGTGGGTTCTTCTGTGAAGAGAATTCAGTGT
GATGATGTGATTTCTGAAAGAAGAAATTCAGAGGGGTCAATATCAAGCAGGGATGTTTACTGAAGCAGGGGCATAGAAGGAAAAAC
20 TGGAAAGTGAGGAAGTTTCACTCTTGAGAGAAGACCCGTGCCCTACCTGCACTACTATGACCCGTCTGGGGCAGAAGATCCCCTGGGA
GCAATTCACCTTGAGAGGCTGTGTGGTGACTTCAGTGGAGAGCAACTCAAATGGCAGGAAGAGTGAGGAAGAGAACCTTTTGTAG
ATCATCACAGCAGATGAAGTGCACTATTTCTTGCAAGCAGCCACCCCAAGGAGCGCACAGAGTGGATCAAAGCCATCCAGATG
GCCTCCCGAAGTGGGAAGTAAAGAGACTCCTGCATTCTCTCTCCCTCTCTGAGGGAAGCCCATGGACAAGCTCAGTCCAGGACC
TGTCACCTTCTGTGACAAATCAACGGGAAACAGCCAGGGGTGGGAAGTTTTCATTGTCAGGGGGGTCTGAATGTAACTCACC
25 TGTGGTGTGAAGGTTCCTCTGCATTGTATTGCTCACTGCAGCCCTCTGCCCTATCCATGACCCCAAGCAGATATAACAAG
CTGTGCAGCCTCAGTAGGCTGCTTCCCTCTCCAGCCTCAGGGCCTCTTCTGAAAAATGAAGAAATTCAACTAGTAGATTCTCTG
AGGTCCCCCTAGCTTAAAAAATAAATACTGCCCATGATTCTTAACACTCGCAGTAGTGATAGTATCTAGTTGTTCTGCTG
GTGTCTCTCTTGGCTAAGTCTTGGCCTTCACTTCTTCAAAATGTACCAGAACCTGAGCCAACGCCCTCCCTGTGAAAGTCTGTG
CTGATCTGTAGTACAGTACCAGGAAGAAACCTCTTTTGTCTCTTTAGACATCTTCTACTTGTCTTGGCCTTGAGATCGTGTA
30 ACAAATGAAGGAGGCTCTCTTCTTCTTCTCTCACTCAAAAACTTCCCGAGAGCAGTGGTGGTTTGTAGGGTTTGTGAC
TTCTATTACTTTTGGCAGCCTGGAAAGTTGTGTCTTCTGGGAAAGAGACCCGGGGAGGCCAGGAGTAGCTGAGGGTCTTTCTG
TGCCCTTAAACCGCCAGAGGAGCCCTATTCCACTCTGGTCTTAGGCTGATCTGAGAGGGTCTCCCTTTGTCTCTTCTGGAGC
ATTTCTCTAACGTTTATTACAATTAGGAGGGGGACCCACATCTGTGAGATTCTGTTTCATTTGAGGTTTACAGAAAAA
GTGGCCAGATGTGTTCCCCCATGGGTGAGAGGCCCTGGGCAACTGCCTGGTGAATGTGTCTTGGCCAGCTGCAGCAAGTGGAG
35 GGGCTGAACACTGCGCCAGCTCACTGGATGATGGGTTAATACAACACTGCAGTGAAGGACTCAGAGCCACAGAACTTCTG
AGAGGGGCTGTAGCATTGCGCAGCATCTTCACTTCTCCAGTAAATGATATGCGTTTCTGTCCTCAGCTTTAAGCACAAGTAGC
AGCAGCTCTGCTTGAGTTCTGAGGGCATCATGGCCCTATGATTAAACAGAGTGATCTAACCTAGACTAAATTTGGAACCTTAT
TTGCAATTTTGTACCTGACCACTAACTAGTGATTCTTCTCCAAATTTGAGAAAGACAGCACCATTGAAACAGATATGTGTGT
GAAAGTATATTTTCAATTCAGATTTTAAATTTTAAAGGCTCCAGGAAAGAAAGGAGAGTAGAACATTTTCTCTCATTTTATCA
40 AATCCTCTCTTGGCCCTCCCTCAATTCCTCTGAAGCTGTTCCTACTCCAGATGGTTTTATCAATAGCCTAGAG
GTAAAGAACTGTCTTTTCTCTGATTCTTTAATAAATTATCTTTATAGAATATGCACAAGTTTTCTTACTACTAGTGTAAAGT
ATTTATTAATGGGAAGTCAACTTAATGTTTTGAAATAAATATATGACTCTGTTTAAAT

433

MLLQSQTMGVSHSFTPKGITIPQREKPGHMYQNEQDYLQNLPTETTVLGTVQILCCLLISSLGAILVFAPYPHFNPASTTLM
45 SGYPFLGALCFGITGSLSIISGKQSTKPFDLSSLTSSNAVSSVTAGAGLFLADSMVALRTASQHCSEMIDYSSLPYSEYYYPI
YEIKDCLLTSVSLTGVLVVMLIFTVLELLLAAYSSVFWKQLYSNNPGSSFSSTQSQDHIQVKKSSRSWI

434

- AAGTCTCAGAGGCTGGAGAGCAGAGCACCAGATCGTTCTGGCAGGAACAGCCAGTGGGAGGTTCCAGCTGAGCGCTCCCCAGA
GGTGAGCTGATCCCCAGCCACAGCACAGGACCAGGCTGCGAGAACAGCATCATCAGCATCATGCTATTACAATCCCAAACCA
TGGGGGTTTCTCAGAGCTTTACACCAAAGGGCATCACTATCCCTCAAAGAGAGAAACCTGGACACATGTACCAAACGAAGATT
ACCTGCAGAACGGGCTGCCAACAGAAACACCGTTCTTGGGACAGTCCAGATCCTGTGTTGCCTGTGATTTCAGTCTGGGGG
5 CCATCTTGGTTTTGTGTCCTTACCCCTCCCACTTCAATCCAGCAATTTCCACCACCTTTGATGTCTGGGTACCCATTTTAGGAG
CTCTGTGTTTGGCATTTACTGGATCCCTCTCAATTATCTCTGGAAAAACAATCAACTAAGCCCTTTGACCTGAGCAGCTTGACCT
CAAATGCAGTGAGTTCTGTACTGTCAGGAGCAGGCCTCTTCTCTCTGCTGACAGCATGGTAGCCCTGAGGACTGCCTCTCAAC
ATTGTGGCTCAGAAATGGATTATCTATCCTCATTGCTTATTCGGAGTACTATTATCCAATATATGAAATCAAAGATTGTCTCC
TGACCAGTGTCAAGTTTAAACAGGTGTCTAGTGGTGTGCTCATCTTCACTGTGCTGGAGCTCTTATTAGCTGCATACAGTTCTG
10 TCTTTTGGTGGAAACAGCTCTACTCCAACAACCTGGGAGTTTATTTCTCGACCCAGTCACAAGATCATATCCAACAGGTCA
AAAAGAGTTCTTACGGTCTTGGATATAAGTAACCTTGGCCTCAGAGGAAGGAAAAGCAACTCAACACTCATGGTCAAGTGTG
ATTAGACTTTCTGAAATCTCTGCCATTTTAGATACTGTGAAACAACTAAAAAAGCTTTTGTGTTGTATTGTTTAC
TATGAGTCGTATTATTTCTCTTGAATAATTTCTTCAAAGCCCAAGTCAATAAA
435
- 15 MVNVPKTRRTFCKKCGKHQPHKVTQYKKGKDSLVAQRRRYDRKQSGYGGQTKPIFRKKAKTTKKIVLRLECEPNCRSKRMLA
IKRCKHFELGGDKRKGQVIQF
436
- GGCGAGAGCTGCGAAAGGCGAGAGCTGCGAAGGGCCAGGTGTGGGCGCTGTTTCTCGTTTTCATCATATAGACAAAAAGCCCC
TGCTGCAAGATGGTCAACGTACCTAAAACCCGAAGAACCTTCTGTAAGAAGTGTGGCAAGCATCAGCCTCACAAGTGACACA
20 GTATAAGAAGGGCAAGGATTTCTTGTATGCCAGGGAAGGAGGCGCTATGATCGGAAGCAGAGTGGCTATGGTGGGCAGACAAA
GCCAATTTTCCGGAAGAAGGCTAAGACCACAAAGAAGATTGTGCTAAGGCTGGAATGTGTTGAGCCTAAGTGCAGATCCAAGAG
GATGCTGGCCATTAAAGAGATGCAAGCATTTTGAAGTGGGAGGAGATAAGAAGAGAAAGGGCCAAGTGATCCAGTTCTAAACTTT
GGGATATTTTCTTCAATTTTGAAGAGAAAAATGGTGAAGCCATAGAAAAAGTTACCCGAGGGAAAAATAATACAGTGATATCTT
ACGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
437
- 25 MATNGSKVADGQISTEVSEAPVANDKPKTLVVKVQKKAADLPDRDTWKGRPDLMSCVGYAIGLGNVWRFPYLCGKNGGGAFLI
PYFLTLIFAGVPLFLECSLQYTSIGGLGVWKLAPMFKGVLAAAVLSFWLNIYIYIISWAIYYLYNSFTTLPWKQCDNFW
NTDRCSFNSYSMVNTTNTMTSAVVEFWERNMHQMTDGLDKPGQIRWPLAITLAIWILVYFCIWKGVGWTKGVVYFSATYPYIMLI
ILFFRGVTLPAGEGILFYITPNFRKLSDESEVWLDAAQTQIFFSYGLGLSLIALGYSNPFHNNVYRDSIIVCCINSCTSMFAGF
30 VIFSIVGFMAHVTKRSIADVAASGPGLAFLAYPEAVTQLPISPLWAILFFSMLLMLGIDSQFCTVEGFITALVDEYPRLLRNRR
ELFIAAVCIISYILGSLNITQGGIYVFKLFDYYSASGMSLFLVFFECVSIWIFYGVNRFYDNIQEMVGSRPCIWNKLCWSFFT
PIIVAGVFIFSAVQMTPLTMGNYPVPKWGQGVGWLMLSSMVLIPGYMAYMFLALKGLKQRIQVMVQPS EDTVRPENGPEHAQ
AGSSTSKEYI
438
- 35 GAATTCGCTCCGCGCCGAGGATCTCCCAAGGTGGCAGAAGGAGGCTTCTGGAGCTGACCCACCCCGACGACCATCAGGGT
GAGGCAACTCCAAGGTCTACTCTCTTTCTGTGCTGTATTACCCACCCCGTCTCTAGGGTGCCCTTGAGCCGCAAACTGCTG
TCCACGTGGACCGGGGTGACATCGCACGTCCATCTGCCAGGACCCCTGCGTCCAAATTCGAGACATGGCGACCAACGGCAGC
AAGGTGGCCGACGGGCAGATCTCCACCGAGGTCAGCGAGGCCCCGTGGCCAATGACAAGCCCAAAACCTTGGTGGTCAAGGTG
CAGAAGAAGCGGCAGACCTCCCGACCGGGACACGTGGAAGGGCCGCTTCGACTTCTCATGTCTGTGTTGGGCTATGCCATC
40 GGCCTGGGCAACGTCTGGAGGTTCCCTATCTCTGCGGGAAAAATGGTGGGGGAGCCTTCTGATCCCTATTTCCTGACACTC
ATCTTTGCGGGGGTCCCACTTCTCTGTGAGTGTCTCCCTGGGCCAGTACACCTCCATCGGGGGCTAGGGGTATGGAAGCTG
GCTCTATGTTCAAGGGCGTGGGCTTGGCGCTGCTGTGCTATCATTTCTGGCTGAACATCTACTACATCGTCATCATCTCTGG
GCCATTTACTACCTGTACAACTCCTTACCACGACACTGCCGTGGAACAGTGCAGACAACCCCTGGAACACAGACCGCTGCTTC
TCCAACTACAGCATGGTCAACACTACCAACATGACCAGCGCTGTGGTGGAGTCTTGGGAGCGCAACATGCATCAGATGACGGAC
45 GGGCTGGATAAGCCAGGTGAGATCCGCTGGCCACTGGCCATCACGCTGGCCATCGCCTGGATCCTTGTGTATTTCGTATCTGG
AAGGTGTTGGCTGGACTGGAAAGGTGGTCTACTTTTACGCCACATACCCCTACATCATGTGATCATCTCTTCTCCGTGGA
GTGACGCTGCCCGGGCCAAGGAGGGCATCTCTTCTACATCACACCCAACCTTCCGCAAGCTGTCTGACTCCGAGGTGTGGCTG

GATGCGGCAACCCAGATCTTCTTCTCATACGGGCTGGGCTGGGGTCCCTGATCGCTCTCGGGAGCTACAACCTCTTTCCACAAC
AATGTCTACAGGACTCCATCATCGTCTGCTGCATCAATTCGTGCACCAGCATGTTTCGAGGATTTCGTATCTTCTCCATCGTG
GGCTTACATGGCCCATGTACCAAGAGGTCATTTGCTGATGTGGCCGCTCAGGCCCCGGGCTGGCGTTCCCTGGCATACCCAGAG
GCGGTGACCCAGCTGCCTATCTCCCCACTCTGGGCCATCCTCTTCTTCTCCATGCTGTTGATGCTGGGCATTGACAGCCAGTTC
5 TGCACTGTGGAGGGCTTCATCACAGCCCTGGTGGATGAGTACCCAGGCTCCTCCGCAACCCGAGAGAGCTCTTCATTGCTGCT
GTCTGCATCATCTCCTACCTGATCGGTCTCTCTAACATCACTCAGGGGGGATTTATGTCTTCAAACCTCTTGTACTACTACTCT
GCCAGTGGCATGAGCCTGCTGTCTCGTGTCTTTGAATGTGTCTCTATTTCTGGTTTTACGGTGTCAACCGATTCTATGAC
AATATCCAAGAGATGGTTGGATCCAGGCCCTGCATCTGGTGGAACTCTGCTGGTCTTTCTTACACCAATCATTGTGGCGGGC
GTGTTCAATTTTCAGTGTGTGAGATGACGCCACTCACCATGGGAACTATGTTTCCCAAGTGGGGCCAGGGTGTGGGCTGG
10 CTGATGGCTCTGCTTCCATGGTCTCATCCCCGGGTACATGGCCTACATGTCTCGCCCTAAAGGGCTCCTGAAGCAGCGC
ATCCAAGTCATGGTCCAGCCCAGCAAGACACTGTTCCGCCAGAGAAATGGTCTGAGCAGCCCAGGCGGGCAGCTCCACCAGC
AAGGAGGCCTACATCTAGGGTGGGGCCACTACCCAGCCGACACTCTACCCCCGACCTGGCTGAGTGCACCACTTGA
TGCTTGAGATACCTTCCATCTCAACCTACCTCGAGTGGTGATCCAGACACCATCACCACGAGAGGGGAGGTGGGAGGACA
GTTAGACCCCTGGGTGGGCCCTGCGGTGGGCAAGGATACCCGGTGGCTTCTGGCACTGGCGGGCTGGTGACCTTTTTAATCCAG
15 GCCCCATCAGCATCCCACTCCTGGCGGGAT

439

MEGYSEEASLLRHLEKVAEEEEVPLVVYLKENAALLTANGLHLSQNRQAQSSPAPPPAEVHSPAADVQNQLASPSATLTTPT
SNSSHNPPATDVQNPPATVVPQSLPLSSIQQNSSEAQLPSNGTGPASKPSTLCADGQQAPEEVRCSLLIDKVSTPATTTTS
TFSREATLIPSSRPASDFMSSSLIDIQNTLVVSADQEMSGRAAATPTKVYSEVHFTLAKPPSVNRTARPFQIAPGGTS
20 QMERSPMLERRHFGKAPAPQPPSLPDRSPRPQRHIMSRSPMVERRMMQORSPASERRPLGNFTAPPTYTETLSTAPLASWVRS
PPYSVLYPSSDPKSSHLKGQAVPASKTGILEESMARRGSRKSMFTFVEKPKVTPNPDLDDLVTADKRRQRDQGEVGVVEEBE
FALGAEASNFFQEPAPRDRASPAAAEVPEWASCLKSPRIQAKPKPKPNQNLSEASGKAELYARRQSRMEKYVIESSSHTPE
LARCPSPTMSLPSSWKYPTNAPGAFRVASRSPARTPPASLYHGYLPENGVLRLPEPTKQPPYQLRPSLFLVSPIKEPAKVSRAA
SPAKPSSLDLVPNLPGKALPPSPALPRPSRSPGLYTSFGQDSLQPTAVSPPYGGDISPVSPSRAWSRAKQAPRPSFSTRNAG
25 IEAQDRRESLPTSPWTPGASRPSSLDGWVSPGPWEPRGSSMSSPPPLPPPPMSPSWSERSVSPRLPETEARPPSRQLQAL
LARNIINAARRKSASPRSAGAENPRPFPSPRAPPPPPPPPPMRSPQPARPGSAAVPGAAPAPIPRSPLPAGPSSCTSPRS
PLPAPPRPFLYRRSPTDSVSLDSEDGAKSPGILGYNICPRGWNGLRLKRGSLPAEASCTT

440

GAGAAAAGTCACATCCAGCTCCTTCATGTGTGCTGCCGATGTGGGATTTTCTGGCTTCGTCAGCCAAGCAATTGGCTGCCTTTG
30 CCGGAGGCTTTGAGAACCAAGGCTTGAAGGCCGCGCAGGAGCATCCAGGGGGAAGCTTGGCTTCAGGGAGGACCTAGCAGACGTT
GGGCGGAGCACTAGCCTCAGGAGAAGGATCTGAAAGAAGCAAGGCGCGGAGCCAGCAGATTGCAGCCCAGCTGACCACCCC
TCCCAGCTCCAATTCCCGTGGCGTCCAGCTCTTCAACAGCGCGCGGAGAGGGTGAACGAGTTACCTTGGAGAGCCACGGCCA
GAGGGGACAGAAGCCCAGCCAGGAGTCCCTCAGAGTGTCTTCAAGCCTCCCAGGCCATGCACCGGGGCTCAGCTGAGTTT
CACCTCGCTGCCGAGCCAGGCCCTCCACGGCACCCAGTCCCAGAGCCCCGACAGAGGGTCCCTGGCCACAGCATGGAGGG
35 GTACTCAGAGGAGGCTAGCTTGCTCGCGCACCTGGAGAAGGTGGCCAGTGAGGAGGAAGAGGTACCACTGGTGGTTTATCTAAA
GGAGAATGCAGCACTGCTGACAGCCAATGGGCTGCACCTGTCCCAAACCGAGAGGCCAGCAGTCCCTACCGGCCCCACCTCC
AGCTGAGGTCCACAGCCCAGCTGCAGATGTCAACCAAAACCTTGCTCGCCAGTGCCACGCTCACCACACCAACTTCTAACAG
CAGCCACAATCCGCCAGCCACCGATGTCAATCAGAACCACCGGCAACTGTTGTCCACAGAGCCTGCCACTTTCTAGCATCCA
ACAGAATTCCTCAGAGGCCAACTCCCATCTAATGGCACAGGGCTGCTTCCAAACCCAGCACCTGTGTGCTGATGGGCAACC
40 CCAGGCACCGGCTGAGGAGGTGAGATGCAGCACACTCTAATGACAAGGTATCAATCCAGCTACCAACCAGCACTTCTC
CAGAGAAGCTACGCTCATCCCCAGCTCCAGGCCCCAGCCTCAGATTTATGTCAGCTCCCTGCTCATTGACATCCAGCCAA
CACCTAGTGGTGTGTCAGCAGATCAAGAGATGTCTGGCGGAGCAGCTGCCACCACGCCCAAGGTCTACAGTGAGGTCCACTT
CAGACTGGCCAAAGCCCCATCAGTGTCTAACAGGACGCCAGGCTTTTGGGATCCAGGCGCCAGGGGGCACCAGCCAGATGGA
GAGGAGCCCCATGTAGAGAGACGACATTTGGGGAGAAGGCCCGGGCTCCCCAGCCCCCAGTTTGGCAGACAGAGACCCCCG
45 GCCACAGAGACATAATGTCCCGCAGCCCCATGGTGGAAAGGAGGATGATGGGGCAGCGAAGCCCGGCTCAGAGAGACGCC
CTTGGGGAACCTTCACTGCACCCCCACCTACACTGAGACCTTGTCACAGCCCCCTTGCTTCTGGGTGAGGTCTCTCTCCCTC
ATATTCTGTCTGTATCCAGCTCCGACCCCAAGTCTTCTCATCTGAAGGGCCAGGCGGTTCTTGCAGCAAGACGGGCACTT
GGAGGAGTCGATGGCCCGCGGGGAGCCGCAATTCATGTTTACTTTCTGTTGAGAAGCCCAAGGTGACCCCAATCCAGACTT
GCTGGATCTGGTACAGACAGCGGATGAGAAGCGGCGGAGAGGACCAGGGGGAGGTAGGCGTGGAGGAGGAGCCCTTCGCACT
50 GGGGGCCGAGGCTCCAATTCAGCAGGAGCCAGCACCTCTGTACAGGGCCAGCCCCGCGGCGGAGGAGGTGGTACCAGA

[illegible]

441

45 MFGPHKPKMYRSIEGCCICRAKSSSRFTDSKRYEKDFQSCFGLHETRSGDICNACVLLVKRWKLPAGSKKNWNHVV DARAGP
SLKTTLPKPKVKVTLSGNRIKSNQISKLQKEFKRHNSDAHSTTSSASPAQSPCYSNQSDGSDTEMASGSNRTPVFSFLDLTYWK
ROKICCGIYYKGRFGEVLIDTHLFKPPCCSNKKA AEKPEEOGPEPLPISTOEW

442

AGACTATCTTTCTAGACAAGGCAGTTGAGGAGGAGGGAGCGCTTGAGGGGGA CTGGCTGGCGTGCACTCCGCACCTCGGGGAC
ATTATTGCGCGTGGAACGGCTGCTTTTGAAGGCACAAC TTCTGAATGGACCATGACTCCACCAAAGATCCC GTCTCTGAT
50 TCACCAAACAGCTTCAACCTGAAACCAGGACGAGAAGTTGACAACATCTGAGTGGACAGCTAAATTGACCTAAGACTTCAGACC

AGACTATTGCCAGAGAAAGATGTTTGGTTTTTCAAGCCAAAGATGTACCGAAGTATAGAGGGCTGCTGTATTTGCAGAGC
TAAGTCCCTCCAGTTCTCGATTCTGACAGTAAACGCTATGAAAAGGACTTCCAGAGCTGTTTTGGATTGCATGAGACTCGTTTC
AGGAGACATCTGCAATGCCCTGTGCTCTGCTGTGAAAAGATGGAAGAAGTTGCCAGCAGGATCAAAAAAACTGGAATCATGT
GGTAGATGCAAGGGCTGGACCCAGTCTAAAGACTACATTTGAAACCAAAGAAAGTGAAGAACTCTATCTGGGAACAGGATAAAAG
5 CAACCAGATCAGTAAACTGCAGAAGGAATTTAAACGTCATAATTTCTGATGCTCACAGTACCACCTCAAGTGCCTCCCCAGCTCA
ATCTCCTTGTACAGTAACCAGTCAGATGACGGCTCAGATACAGAGATGGCTTCTGGTTCTAACAGAACACCAGTTTTTTCCTT
TTTAGATCTCACTTACTGGAAAAGACAGAAGATATGTTGTGGGATCATCTATAAAGGCCGTTTTTGGGGAAGTCCCTATTGACAC
ACATCTCTTCAAGCCTTGCTGCAGCAATAAGAAAGCAGCTGCTGAGAAGCCAGAGGAGCAGGGGCCAGAGCCTCTGCCCATCTC
CACTCAGGAGTGGTGAAGTGTGAGTTTATGTAGAGGGGAACAAAAAAATCTGAATTTTGAAGAACCAAGCTACAAA
10 CTGACCCCTCTTTTTTTTTTTTTTTTTTGTAGAGCGGATTTGCTCTTGTATTACCCAGGCTGGAGTGCAGTGGCGTATCTTGGCT
CACTGCAACTTCCGCTCTCCCGGGTTCAAGTGATTTCTCTGCTCAGCCTCCCAAGTAGCTGGGTTTATAGTGGCCCGCCACCAG
ACCCGGCTAATTTTTTTAGTTTTAGTAGAGACGGGGTTTACCACGTTGGCCAGGCTGGTCTTAAATGACCTCTTTATTTTTAAC
TTGGATACCTGCTATTCTGCCAAAAGACAATTTCTAGAGTAGTTTTGAATGGGTGATTTCCTCCCACTCCCAAACTCTGAAG
CCAGTGTCTAGCTTACTAAAAAAGAGTTGTATATAATTTAAGATGCTGAGTATTTATAGGAAAGCTGAATGCTGCTGTAA
15 AGTGCTCTTTAAGCTTTTTTTTTTTTAAATCCCTTCTAATGAATGAAACTAGGGGAATTTACAGGGACAGAGATGGGATTTGTT
GTATGATAAACTGTATGTAGTTTTTAGTCTTTCTGTTTTGAGAAGCAGTGGTTGGGGCATTTTTAAGATGGCTGGCTACTCTTG
TTTTCCCTCATGATAATAAATTTGTCATAACTCAGTAACATGAACCTGCCCCCTAGAGGTAGTTGTTAATAATTTTGAATATTA
AGGTCTTGCCAAGCTTCTGATGATTCACACCTGTACTACTGATTATTAAGCAGGACAGACTGAGCTTTCTGTTGCAATACCTT
GGAGGAGAAAGTAATTTCTAAATATACAGAGAGGTAACCTGACTATATATGTTGCATCCTGTGCTCCCTCATATTAATATTT
20 GATAAAGATTTTAAATTTATGTAAACTTCTAAAGCAGAACTCAAGCTCCTCTTGGGGAATGGCAAGCTTTAGGATAGGCAAG
ACCCGTGTATGAATAGTACCAAGCATTACCGCATGGTAGAGAACACACTCGATTAAAAATGTTAAGCTATCTGAAAAATAAAAT
GTGCAAGCTTTCAGGATGGCACAACCAAGGTTAATGCTTCTTGGGGCACATTTCTTAGAGGGCTTGCTGAGTGTGTAATAT
AATCGACTTTTGTGTGTGTACATGACTTCTGTGACTTCATTGAAAATCTGCACAATTCAGTTTCAGCTCTGGATTACTTCAGT
TGACCTTTGTGAAGGTTTTTATCTGTGTAGAATGGGTGTTGACTTGTTTTAGCCTATTAAATTTTTATTTTCTTCTACTCTGT
25 ATTAAGTAAACTTACTAAAAGAAAAGAGGTTTGTGTTTACATTAATGTTTGGTTTGGCTTCTTTTAGTCAGGCTTTCT
GAACATTGAGATATCCTGAACCTTAGAGCTCTTCAATCCTAAGATTTTTCATGAAAAGCCTCTCACTTGAACCCAAACAGAGTAC
TCTTACTGCCTCTTTCTAAATGTTTCAAGAAAAGCATTGCCAGTTTCAAGTCTTTTCAAAATGAGGGAGAAACATTTGCTGCTT
GTAATAACAAGACTCAGTGTCTTATTTTTTAACTGCATTTTAAATTTGATAGTATAATAACAATAAGGAGTAAGCCACCTTT
TATAGGCACCCCTGTAGTTTTATAGTTTAAATCTAAACATTTTATATTTTCCCTTCTTTTGGAAAAAACCTACATGCTACAGCCA
30 CCATATGCACAGACTATACAGTGAGTTGAGTTGGCTCTCCACAGTCTTTGAGGTGAATTACAAAAGTCCAGCCATTATCATCC
TCTGAGTATTTGAAATGATTTTTTTTGTACATTTTGGCTGCAGTATTTGGTGGTAGAATATACTATAATATGGATCATCTCTA
CTTCTGTATTTATTTATTTATTACTAGACCTCAACCACAGTCTTCTTTTCCCTTCCACCTCTCTTTGCTGTAGGATGTACT
GTATGTAGTCATGCACTTTGTATTAATATATTAGAAATCTACAGATCTGTTTTGTACTTTTTTACTGTTGGATACTTATAATC
AAAATTTTACTAGGGTATTGAATAAATCTAGTCTTACTAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA
35 443
MKSQWCRPVMADLGVYQLRHFSISFLSSLLGTENASVRLDNSSSGASVVAIDNKIEQAMDVLKSHLMYAVREBEVLKEQIKEL
IEKNSQLEQENLLKTLASPEQLAQFQAQLQTGSPPTTQPPQGTQPPAQPASQSGSPTA
444
CGCCTCTTACGGCACTGGGATCCGCATCTGCCCTGGGATCATCAAGCCCTAGAAGCTGGGTTTCTTTAAATTAGGGCTGCCGTT
40 TTCTGTTTTCTCCCTGGGCTGCGGAAAGCCAGAAGATTTATCTAGCTTATACAAGGCTGCTGGTGTTCCTCTTTTTTCCACG
AGGGTGTTTTTGGCTGCAATTGCATGAAATCCCAATGGTGTAGACCAGTGGCGATGGATCTAGGAGTTTACCACTGAGACATT
TTTTCAATTCTTTCTTGTATCTCTGCTGGGACTGAAAACGCTTCTGTGAGACTTGATAATAGCTCCTCTGGTGCAAGTGTGG
TAGCTATTGACAACAAAATCGAGCAAGCTATGGATCTAGTGAAGGCCATTTGATGTATGCGGTGAGAGAGAGTGGAGGTC
TCAAAGAGCAAAATCAAAGAACTAATAGAGAAAAATTTCCAGCTGGAGCAGGAGAACAACTGCTGAGAACACTGGCCAGTCTCTG
45 AGCAGCTTGGCCAGTTTTCAGGCCAGCTGCAGACTGGCTCCCCCCTGCCACCACCCAGCCACAGGGCACACAGCCCCCG
CCCAGCCAGCATCGCAGGGCTCAGGACCAACCGCATAGCTGCCATGCCCCCGCAGAACTGGCTGCTGCGTGTGAACGTGAACAG
ACGAGAGAAGATGTGCTAGGGAGAACTGCCTCCACAGTCAACCATTTTATGCTCGCTGCGAAAGAGACGTGAGACTGACATAT
GCCATTATCTCTTTTCCAGTATTAAACACTCATATGCTTATGGCTTGGAGAAATTTCTTAGTTGGGTGAATTAAGGTTAATCC
GAGAATTAGCATGGATATACCGGGACCTCATGCAGCTTGGCAGATATCTGAGAAATGGTTTAAATTCATGCTCAGGAGCTGTGTG
50 CCTTCCATCCCTTCCGGCTCCCTACCCCTCACTTCCAAGGTTCTCTCTCTGCTTGGCTTAGTGTCTACATGGGGTTGTG

[illegible]

GGGCTGCGATGGACATTTCGGAATTTCTTTGGAGTAATACCAAGTGGAAAGAACTTGTAAAGTAAACAGTAAAGAAGAATGAGA
AAACAAAGTCTGATGAAGAACTTTAAAAGCAAAGAAAGGAATAAGGAAATCAAGGTAATAGCTCCCGTAAAGAGGATGACT
TCAACAAAAGCAACCAAGCAAGAAAAAGAGGATCATCTATGATTCAGATTCAGAGTCAGAGGAGACGTTGCAGGTAAAAAATG
5 CCAAAAAGCCACCAGAAAACTGCCAGTATCTTCTAAACCTGGTAAAAATTTACGGCAGGATCCTGTTACATACATTTTCAGAAA
CAGATGAAGAAGATGACTTTATGTGTAAGAAGCGGCCCTCTAAATCAAAAGAGAATGGAAGATCTACAAATAGTCATCTTGGAA
CATCAAAACATGAAAAAGATGAAGAAACACTAAGACCAAGAATAAGCCTTTATCACCATAAACTTACACCCACATCAGTAC
TTGATTATTTTGGAACTGGAAGTGTCCAAAGATCTAATAAGAAGATGGTGGCAAGCAAAAGAAAAGAGCTTTTACAAAAATACAG
ATGAGTCTGGATTAAATGATGAAGCCATCGCCAAGCAATTCAGCTTGATGAAGATGCGGAGCTGGAGAGGCAGTTGCATGAAG
ATGAAGAGTTTGGCAGAACATTAGCCATGTTGGATGAAGAACCACAGCAAAAAGGCTCGAAAGGACACAGAAGCGGGAGAAA
10 CGTTTTCATCTGTCCAAGCCAAATTAAGTAAAGCAGAAAAACATAAATATCCTCATAAAGTAAAAACAGCACAGTTTCAGATG
AAAGAAAGAGCTACAGTCTTAGGAAGCAAGTAAATGAAAGTTTCAAAAGAATCTCAGCAACATTCCAAGTCATCAGTGTGACA
AAATAGGAGAGTCTCTTCTTCCCAAGGCCAGTTCTAAGCTGGCAATTTATGAAAAGAAAAGAGAGCTCTTATAAAGAAATAG
AGCTGTGGCTTCAAAAAGAAAAGAAAATGCCATTAATTTGAAAGGAGAGACAAAAACTCTAAGAAAACAAAAGTTCTCCAG
CTAAAAAGAGTCTGTAAAGTCTGAAGATCTGAAAAGAAACGCACTAATTATCAAGCTTATCGAAGCTACTTAAATCGAGAAG
15 GTCCCAAGGCTCTGGGCTCCAAAGAAATACCAAGGGAGCTGAAAATTTGCTTGGAAAGGCCTTATATTTGTAATCACAGGCGTGC
TGGAGTCTATTGAACGAGATGAGGCCAAGTCTCTAATTGAACGTTATGGGGGAAAAGTAACAGGAAATGTCAGCAAGAAAACAA
ATTATCTTGTCTATGGGTCGTGATAGTGGACAGTCCAAGAGTGATAAGGCCGAGCCTTGGGGACAAAAATTTATGATGAAGATG
GCCTGTTGAATCTGATTCCGACTATGCCAGGCAAGAAATCCAAGTATGAAATAGCAGTTGAAACTGAGATGAAGAAAGAGTCCA
AACTGGAGAGAACACCCCAAAAAAATGTCCAAGGAAAAAGAAAATTAGTCCATCTAAAAAGGAATCAGAATCTAAAAAGAGCA
20 GGGCGACTTCCAAAAGGGACAGTTTGGCAAAGACAATAAAAAAGGAAACAGATGTGTTTTGAAAAGCCTGGATTTCAGGAGC
AGGTGGCTGAGGAGACAAGTGGTGACAGCAAGGCTAGGAATTTGGCTGATGACAGCAGTGAACAAAGTGGAAAATTTGCTCT
GGGTGGATAAATATAAGCCAACCTCGCTCAAGACCATAATTGGACAGCAAGGTGACCAGAGCTGTGCCAACAACTCCTACGCT
GGCTCCGAAACTGGCAAAAGAGTTCTTCCGAAGATAAAAAACCGCAGCAAGTTTGGTAAATTTTCCGGCAAGATGATGGCT
CTAGTTTTAAAGCAGCGTTGCTGTGTCAGGCCCTCCTGGTGTGGCAAAACCACACAGCTTCCCTGGTGTGTCAGGAGTTGGGAT
25 ACAGCTACGTGGAAGTGAATGCAAGTGAACCCGGAGTAAGAGCAGTTTGAAGGCGATTGTTGCTGAGTCACTGAACAATACCA
GCATCAAAGGCTTTTATTCAAATGGAGCAGCCTCTTCAGTAAGCACGAAACATGCTCTCATCATGGATGAAGTAGATGGCATGG
CAGGCAATGAGGATAGGGGAGGAATTCAGGAATTAATTGGCCTGATAAAACATACTAAAAATCCCATTAATTTGTATGTGCAATG
ATAGAAATCATCCCAAGATTCGCTCTCTGGTTCATTATGTTTTGATCTTCGTTTTCAAGACCTCGGGTTGAACAGATTAAGG
GTGCTATGATGTCATTGTCATTTAAAGAAGGTTTAAAGATTCCTCCCTCCAGCTATGAATGAAATAATTTGGGAGCCAAATCAAG
30 ATATCAGACAGGTTTACATAAATCGAGTATGTGTGTCAGCAAGTAAAGCATTAACCTATGACCAGGCCAAAGCTGATTCTC
ACAGAGCCAAAAAGGATATCAAAATGGGCCCATTTGATGTTGCCGAAAAGTGTTTGACGCTGGAGAGGAGACTGCTCAGATGT
CACTTGTGGACAAGTCAGATCTCTTTTTTCATGATTATTCAATAGCACCCCTCTTCGTCCAGGAAAATTACATACAGTGAAGC
CTGTAGCAGCAGGGGTGACATGAAAAGCACCTGATGCTTTTAAAGCAGAGCAGCAGCAGCATATGCGATGGTGACCTAGTGG
ACAGCCAGATCCGGAGTAAGCAAACTGGAGTCTTCTGCTGCGCAGGCCATTTATGCCAGTGTCTTCTCGGAGAGTTGATGA
35 GGGGTACATGACCCAGTTTCCACCTTCCCAAGCTGGCTGGGGAAGCACTCGTCTACAGGCAACATGATCGTATTGTTTCAGG
ACCTGGCCTTGCAATATGAGTCTCAGAACTTACTCCAGCAAAAGGACTGTAAACATGGATTATCTGTCGCTTCTAAGGGATGCAC
TTGTACAGCCCTTGACCTCACAAGGAGTAGACGGAGTACAGGATGTTGTTGCACCTTATGGACACATATTATTTGATGAAGAAG
ACTTTGAGAATATCATGGAATCAGCAGCTGGGGTGGCAACCTAGTCCCTTTTCAAGCTGGATCCCAAGGTGAAAGCAGCCT
TCACAAGAGCTTACAATAAGGAAGCCACCTTACTCCATCTCACTTCAAGCTATATAAGGCATCTAGACACAGCACAAGCCCAT
40 CCCTGGATTTCGGAATACAATGAAGAATTAATGAAGATGACTCTCAATCTGATGAGAAAGACCAAGATGCTATAGAAACTGATG
CCATGATCAAGAAAAAGACAAAATCTTCAAAGCCTTCAAAAACGAAAAAGATAAGGAGCCAGAAAAGGAAAAAGGAAAGTT
CGAAGAAATGAAACCATTTTTCATAGCGACAGCCACTTTTACTCTCCCTCCTGACCAGTCCAGCTGGTCTAGAGAAAGCCTT
GTTTTTCCAGAGCAACCATGTTTTAGCATAATGGGATGACCTGGTGTCCCATTAATAAAGGGTGGTATGGCTAGAAGGG
TATGAGCAGTAGGCTTATGTACACCTCTTATAGAGTTGATAGGACTGCTTGGGCTCTCCACTGTCTCTGTCAATCTAGTTAG
45 ACGTGTCTCTGAATGACTGTAGAAATGGAAGTAAAGTACACCTGGGCTTTGGAGTCAGATTTTAGTTAAACAATATGAGCCT
GGAGCAGTAGTACTAAGGCGTCTTTGTAGGCTTAAAGATTTATCTTAATGGCTTTTATAGGACCAATGCTGATTTTTTTTAA
GGCAGTTCTTATTTTGGTGAATTTTGTAAATAAATGATTATACAAAACAGTCACCACCTAGAACTGGGTATCTTTGTAC
TTGTGAGATATCTTTCAAAGTATAAATAGGAATAAAAAATGTTCCACATGATATGTGTACAAAATCTTTTAAAGATGTGATCT
TGTAACGAAAAATATTATGTTGGTATATTACTTCTTATTAATTTGCAAAATGATTGGATTAAAAAGCTCACTGTATTCTTAC
50 TAACCTGTTGTGTGAACATAACATTGGCAGTTCAGGAGTCAGAACTCGGATTCAGCCTCACTCCCTCTGTGTGTGTGACC
ACGCTCTGTGTGTGAATAGGATTTTATGCACTGCTGTGTTATGGGGCTTTGCAGCGTGGGGGGTGCAGAAGATTCTCAGACAT
TCTTATCTTCTGGAGTCTCAACCCATCAGGTTTCAGACCAGTAGGAACAGGCTGGGTGAGGCTCTTAATTTCACTACGGTGGG

- GGACAAAATGAACACTGGCTTTTGTTCACCTTTCCTGAGAGATCTAGTCCTGTCTGTGATGGCTTCCACAAACCTGTGG
AAGTTTGGGTTATTCCATATACTGTATGTGTGAGTGGCCGTGACGTATGGACCTAGCCTGGAATTTAGATGTCTAGCTATTTT
TTAATCTTGTGGTAAATGCTCAGAGGTCTGTGTACCTCTGTGAGGTATATACTAAATGAATGTAAATGAAGATCGTATTC
CCTGCTTTAAGGAACGTGTAGTCTTGTTTGGGTGATGAGACAAAGTTCTTGAGACATTTAACTGTATTTACCAGTGTGAAAAA
5 TGTCATTTATGGAGATTAAATTCATTATGGAATAAAACATTGCCAAGCCCTTGCCTGCTG
451
MACAAARSPADQDRFICIYPAYLNNKKTIAEGRRIPI SKAVENPTATEIQDVCSAVGLNVFLEKNKMSREWNRDVQYRGRVRV
QLKQEDGSLCLVQFPSRKSVMLYAAEMIPKLTTRTQKTGGADQSLQQEGESKKKGKGGGG
452
10 CGGAAACTCAGAGCCGGGTTCCCTCCCGGGTTTCTGCCGGGTTTCTCCCTGCCGCTCCTGGGTGTGAGACTCTTGTGAAGATG
GCTTGGCGTGGCGCGCGGTCCCGGCCGACAGGACAGGTTTATTTGTATCTATCCTGCTTATTTAAATAAAGAAGACCATC
GCAGAGGGAAGGCGAATCCCATAGTAAGGCTGTTGAAAATCCTACAGCTACAGAGATTCAAGATGTATGTCAGCAGTTGGA
CTTAACGTATTTCTTGAGAAAAATAAATGTACTCTAGAGAAATGGAATCGTGATGTCCAATACAGAGGCAGAGTCCGGGTCCAG
CTCAACAGGAAGATGGGAGCCTCTGCCTTGTACAGTTCCCATCACGTAAGTCAGTAATGTTGTATGCAGCAGAAATGATACCT
15 AAATAAACAAGGACACAAAAACAGGAGGTGCTGACCAAGTCTTCAACAAGGAGAGGGAAGTAAAAAGGGAAGGAAAG
AAAAAGAAGTAACCTAGTATCAGCATCAAGTATGTGGTACTACTGTAAGAGACATGAATGGAGACTTCTAATTTGTATCGGAGG
GAAACAGAAGCTTTTGTGTCATCATTTAAGTGAAGTGTGAACCTTGTGCTCTCATCTTTATCATCGGAGTTGACAGTGAA
ACAAATTTACATCAGAAGTTTGCATCTCGCGTATATGCCGTATAAAGAATTTTGTCTTTCAATGCAGTTTGTGAAGAA
AATATTTTAAATGGACAATGGACTGTACAATAAGTTACTTGAAATAAGTTGTTCAGATAAATTTCAATTAGATTTAAATAA
20 ACATTATGTCACCTTTTAAGTTAATGAAATAAAATTTGAACTGAAAAA
454
CAAATCCCTCTGGCACAGTGGGATGTGTGGAAAAAAGTGAAGTCAAAGACCTCCTGCGTCGTCTAAGTGGTCAGGCTGAAG
GCAAGCTAATCTCTGTAGAATATGGAACAGAGGAAAAATAAATAACAGTAATATCTGTTTATTCAGGTCACCTCAGAAGTG
GTAGGAACATAGAAAGAGTGTCTTTCTACCTAGGATTTCCATTGAAGGCCCTCTGGCATATGATATAGAAGTAATTTAAGACA
25 ATACATCACCTGTAGTTCAAATATGTTTATTTATATCTTTATGATTTTATTCTCTCTCATTCTCATGGCACTTTTCATAACA
TTATGGCTAACCTCTAATTACAGATTTTGCTTTTGCTTCCCTGAATGAATTACAAGCCTTTTAAAGATATGAAATATGCCCTACC
CGCAGAGCTTGGCACAAGTGGAGTCAATCTTTTAAGTTTAAATATGCATTTTCAGACTCAAATAATTAAGAAGTTTCATTG
ATATCCACTGGTCACATCATACTGTCTATAGGGCAATAAAATCTGTGTAAACTCAAAAAAACA
455
30 MSRRYDSRTTIFSPGRLYQVEYAMEAIGHAGTCLGILANDGVLLAERRNIHKLLDEVFFSEKIYKLNEDMACSVAGITSDAN
VLTNELRLIAQRYLLQYQEPFIPCEQLVTALCDIKQAYTFQGGKRPFGVSLLYIGWDKHYGFQLYQSDPSGNYGGWKATCIGNNS
AAAVSMLKQDYKEGEMTLKSALALAIKVLNKTMDVSKLSAEKVEIATLTRENGKTVIRVLKQKEVEQLIKKHEEEEAEREKK
EKEQKEKDK
456
35 CCGTGGACATCTCAGGTCTTCAGGCTCTCCATCTGGAACATATATAAGTTTCAGAAAACATGTCTCGAAGATATGACTCCAGGA
CCACTATATTTCTCCAGAAGTCCGCTTATACCAAGTTGAATATGCCATGGAAGCTATTGGACATGCAGGCACCTGTTTGGGAA
TTTTAGCAAAATGATGGTGTGTTTGTCTGCAGCAGAGAGACGCAACATCCACAAGCTCTTGATGAAGTCTTTTTTCTGAAAAA
TTTATAAACTCAATGAGGACATGGCTTGCAAGTGTGGCAGGCATAACTTCTGATGCTAATGTTCTGACTAATGAACAAAGCTCA
TTGCTCAAAGGTATTTATTACAGTATCAGGAGCCAAATACCTTGTGAGCAGTTGGTTACAGCACTGTGTGATATCAACAAGCTTT
40 ATACAATTTGAGGAAAAACGTCCTTTGGTGTTCATTGCTGTACATTGGCTGGGATAAGCACTATGGCTTTCAGCTCTATC
AGAGTGACCCTAGTGGAAATTACGGGGATGGAAGGCCACATGCATTGGAAATAATAGCGCTGCAGCTGTGTCAATGTTGAAAC
AAGACTATAAAGAGGAGAAATGACCTTGAAGTCAGCACTTGTCTTAGCTATCAAAGTACTAAATAAGACCATGGATGTTAGTA
AACTCTCTGCTGAAAAAGTGGAAATGCAACACTAACAAGAGAGAATGGAAAGACAGTAATCAGAGTTCTCAAAACAAAAAGAAG
TGGAGCAGTTGATCAAAAAACATGAGGAAGAAGAAGCCAAAGCTGAGCGTGAGAAGAAAGAAAAAGACAGAAAGAAAGGATA
45 AATAGAATCAGAGATTTTATTACTCATTTGGGGCACCATTTCAGTGTAAGAGCAGTCCCTACTCTTCCACACTAGGAAGGCTTTA
CTTTTTTAACTGGTGCAGTGGGAAAAATAGGACATTACATACTGAATGGGTCTTGTCAATTTCTGTCCAATGAATACTTTAT
TGTAACGATGATGGTTACCTTTCATGGACGCTTAACTCTCCACACACATCCCTTTTTTTGGAATAAA

458

45 459

MAPVGVGEKLLLLGPNPAPAAGDLTSEEEEGQSLWSSILSEVSTRARSKLPSGKNILVFGEDGSGKTTLMTKLQGAEHGKKGR
GLEYLVLVSHDEDRDDHTRCNVWILDGDLYHKGLLKFAVSAESLPETLVIFVADMSRPWTVMESLQKQWASVLRHIDKMKIPPE
KMRELERKFVKFDQDYMEPEGCGSPQRRGPLTSGSDEENVALPLGDNVLTHNLGIPVLVVCTKCDAVSVLEKEHDYRDEHLD
FIQSHLRRFCLOYGAALIYTSVKEEKNLDDLYKYIVHKTYGFHFTTPALVVEKDAVPIPAGWNEKKIAILHENFTTVKPEDAY

TAATCAAAATGTTATCCAACCTGTATATTGTTTACTTTATTTGTAAATACTGGTGAACAGTGGTTAATAAATAGTTTATATTCCA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

463

5 MKASAALLCLLLTAAAFSPQGLAQPVGINTSTTCCYRFINKKIPKQRLSEYRRRTSSHCPREAVIFKTKLDKEICADPTQKWVQ
DFMKHLDKKTQTPKL

464

10 AGCAGAGGGGCTGAGACCAAACCAGAAACCTCCAATTCTCATGTGGAAGCCCATGCCCTCACCTCCAACATGAAAGCCTCTGC
AGCACTTCTGTGCTGCTGCTCACAGCAGCTGCTTTTCAAGCCCCAGGGGCTTGTCTAGCCAGTTGGGATTAATACTTCAACTAC
CTGCTGTACAGATTTATCAATAAGAAAATCCCTAAGCAGAGGCTGGAGAGCTACAGAAGGACCACAGTAGCCACTGTCCCCG
GGAAGCTGTAATCTTCAAGACCAAACCTGGACAAGGAGATCTGTGCTGACCCACACAGAAGTGGGTCCAGGACTTTATGAAGCA
CCTGGACAAGAAAACCAAACCTCCAAGCTTTGAACATTCATGACTGAACTAAAAACAAGCCATGACTTGAGAAAACAAATAATT
TGTATACCTGTCTTTCTCAGAGTGGTTCTGAGATTATTTTAATCTAATTTCTAAGGAATATGAGCTTTATGTAATAATGTGAA
TCATGGTTTTCTTAGTAGATTTTAAAGTTATTAATATTTTAATTTAATCTTCCATGGATTTTGGTGGGTTTTGAACATAAG
CCTTGGATGTATATGTCTCTCAGTGTGTAAGAACTGTGGGATGCTCTCCCTTCTCTACCTCATGGGGGTATTGTATAAGTC
15 CTTGCAAGAATCAGTGCAAGATTGTCTTTAATTGTTAAGATATGATGTCCCTATGGAAGCATATTGTTATTATATAATTACAT
ATTTGCATATGTATGACTCCCAAATTTTACATAAAATAGATTTTGTAAAAAA

465

20 MQAHELFRYFRMPPELVDFRQYVRTLPNTLMGFGAFAALTTFWYATRPKPLKPPCDLSMQSVEVAGSGGARRSALLDSDEPLVY
FYDDVTTLYEGFQQRGIQVSNNGPCLGSRKPDQPYEWLSYKQVAELSECIGSALIQKGFKTAPDQFIGIFAQNRPEWVIEQGC
AYSMVIVPLYDTLGNEAITIYNKAELSLVFVDKPEKALLLEGVENKLIPLGLKIIIVMDAYGSELVERGQRCGVEVTSMKAME
DLGRANRRKPKPPAPEDLAVICFTSGTTGNPKGAMVTHRNIVSDCSAFVKATENTVNPDPDTLISFLPLAHMFERVVECVMLC
HGAKIGFFQGDIRLLMDDLKVLQPTVFPVPRLLNRMFDRIFGQANTTLKRWLLDFASKRKEAELRSGIIRNNSLWDLRI FHKV
QSSLGGRVRLMVTGAAPVSATVLTFLRAALGCQFYEGYQTECTAGCCLTMPGDWTAGHVGAAMPNCNLKLVDEEMNYMAAEG
EGEVCVKGPVNFQGYLKDPAKTAEALDKDGLHTGDIGKWLPNGLTKIIDRKKHIFKLAQGEYIAPEKIENIYMRSEPVAVQFV
25 HGESLQAFLIAIVVPDVETLCSWAQKRGFEGSFELCRNKDVKKAILEDVRLGKDSGLKPFQVKGITLHPPELFSIDNGLLTP
TMKAKRPELRNYFRSQIDDLYSTIKV

466

30 TCAACACAGGACAATGCAAGCCCATGAGCTGTTCCGGTATTTTCGAATGCCAGAGCTGGTTGACTTCCGACAGTACGTGCGTAC
TCTTCCGACCAACACGCTTATGGGCTTCGGAGCTTTTGCAGCACTACCACTTCTGGTACGCCACGAGACCAACCCCTGAA
GCCGCCATGCGACCTCTCCATGCAGTCACTGGAAGTGGCGGGTAGTGGTGGTGCACGAAGATCCGCACTACTTGACAGCGACGA
GCCCTTGGTGATTTCTATGATGATGTCACAACATTATACGAAGGTTTCCAGAGGGGAATACAGGTGTCAAATAATGGCCCTTG
TTTAGGCTCTCGGAAACAGACCAACCCCTATGAATGGCTTTCATATAAACAGGTTGCAGAATTGTCCGAGTGCATAGGCTCAGC
ACTGATCCAGAAGGGCTTCAAGACTGCCCCAGATCAGTTCATTGGCATCTTTGCTCAAAATAGACCTGAGTGGGTGATTATTGA
ACAAGGATGCTTTGCTTATTCGATGGTGATCGTTCACCTTATGATACCCCTTGGAATGAAGCCATCAGTACATAGTCAACAA
35 AGCTGAACCTCTCTGCTTTTGTGACAAGCCAGAGAAGGCCAAACTCTTATTAGAGGGGTAGAAAAAAGTTAATACCAGG
CCTTAAATCATAGTTGTATGGATGCCACGGCAGTGAAGTGGTGAACGAGGCCAGAGGTGTGGGGTGAAGTCACCAGCAT
GAAGCGATGGAGGACCTGGGAAGAGCCAACAGACGGAAGCCCAAGCCTCCAGCACCTGAAGATCTTGCAAGTAAATTTGTTTAC
AAGTGGAACTACAGGCAACCCCAAGGAGCAATGGTCACTACCGAAACATAGTGAGCGATTGTTTCACTTTTGTGAAAGCAAC
AGAGAATACAGTCAATCCTTGCCAGATGATACCTTTGATATCTTTCTTGCCTCTCGCCATATGTTTGAGAGAGTTGTAGAGTG
40 TGAATGCTGTGTCATGGAGCTAAAATCGGATTTTCCAAGGAGATATCAGGCTGCTCATGGATGACCTCAAGGTGCTTCAACC
CACTGTCTTCCCGTGGTTCCAAGACTGTGAACCGGATGTTTGACCGAATTTTCGGAACAGCAACACACCGCTGAAGCGATG
GCTCTTGGACTTTGCTTCCAAGAGGAAAGAGCAGAGCTTCGACGCGCATCATCAGAAAACAGCCTGTGGGACCGGCTGAT
CTTCCACAAAGTACAGTCGAGCCTGGGCGGAAGAGTCCGGCTGATGGTGACAGGAGCCGCCCCGGTGTCTGCCACTGTGCTGAC
GTTCTCAGAGCAGCCCTGGGCTGTGCTTTTATGAAGGATACGGACAGACAGAGTGCACCTGCCGGGTGCTGCCTAACCATGCC
45 TGGAGACTGGACCGCAGGCCATGTTGGGGCCCCGATGCCGTGCAATTTGATAAACTTGTGATGTGGAAGAAATGAATTACAT
GGCTGCCGAGGCGAGGCGAGGTGTGTGTAAGGGCCAAATGTTTCAAGGCTACTTGAAGGACCCAGCGAAAACAGCAGA
AGCTTTGGAACAAAGACGGCTGGTTACACACAGGGGACATTTGAAATGGTTACCAAATGGCACCTTGAAATTTACGACCGGAA
AAAGCACATATTTAAGCTGGCACAGGAGAAATACATAGCCCCGAAAGATTGAAATATCTACATGCCAAGTGAGCCTGTTGC

- TCAGGTGTTTGTCCACGGAGAAAGCCTGCAGGCATTTCTCATTTGCAATTGTGGTACCAGATGTTGAGACATTATGTTCCCTGGGC
CCAAAAGAGAGGATTTGAAGGCTCGTTTGAGGAACTGTGCAGAAATAAGGATGTCAAAAAAGCTATCCTCGAAGATATGGTGAG
ACTTGGGAAGGATTTCTGGTCTGAAACCATTTGAACAGGTCAAAGGCATCACATTCGACCCTGAATTATTTCTATCGACAATGG
CCTTCTGACTCCAACAATGAAGGCGAAAAGGCCAGAGCTGCGGAACATTTTCAGGTGCGAGATAGATGACCTCTATTCCACTAT
5 CAAGGTTTGTAGTGTGAAGAAGAAAGCTCAGAGGAAATGGCACAGTTCACAATCTCTTCTCTGCTGATGGCCTTCATGTTGTTA
ATTTTGAATACAGCAAGTGTAGGGAAGGAAGCGTTTCGTGTTGACTTGTCCATTCGGGGTCTTCTCATAGGAATGCTAGAGGA
AACAGAACACCGCCTTACAGTCACCTCATGTTGCAGACCATGTTTATGGTAATACACACTTTCCAAAATGAGCCTTAAAAATTG
TAAAGGGGATACATAAAATGTGCTAAGTTATTTGAGACTTCCTCAGTTTAAAAAGTGGGTTTTAAATCTTCTGTCTCCCTGCTT
TTCTAATCAAGGGGTTAGGACTTTGCTATCTCTGAGATGCTGCTACTTGTCTGCAAATTCCTGCAGCTGTCTGCTGCTCTAAAGA
10 GTACAGTGCCTAGAGGGAAGTGTTCCTTTTAAAAATAAGAACAACCTGCTCTGGCTGGAGAATCTCACAAGCGGACCAGAGATC
TTTTTAAATCCCTGCTACTGTCTCTCTCACAGGCATTCACAGAACCCCTCTGATTCGTAAGGGTTACGAAATCATGTTCTTC
TCCAGTCCCTGTGGTTTTCTGTTGGAGCATAAGGTTTCCAGTAAGCGGGAGGGCAGATCCAACCTCAGAACCATGCAGATAAGGA
GCCTCTGGCAAATGGGTGCTCATCAGAACGCGTGGATTCTCTTTCATGGCAGAAATGCTCTTGGACTCGGTTCTCCAGGCCCTGAT
TCCCGGACTCCATCCTTTTTCAGGGGTATTTTAAAAATCTGCTTAGATTCTATAGTGAAGACAAGCATTTTCAAGAAAGAGTTA
15 CCTGGATCAGCCATGCTCAGCTGTGACGCTGAATAACTGCTACTTTATCTTCACTGAACCACTCACTCTGTGTAAAGGCCAA
CAGATTTTTAATGTGGTTTTTCATATCAAAAGATCATGTTGGGATTAACCTGCGCTTTTCCCCAAAAATAAACTCTCAGGCAAG
CATTTCTTTAAAGCTATTAAGGGAGTATATACTTGAGTACTTATTGAAATGGACAGTAATAAGCAAATGTTCTTATAATGCTAC
CTGATTTCTATGAAATGTGTTTGACAAGCCAAAATTCAGGATGTAGAAATCTGGAAAGTTCATTTCTGGGATTCACCTTCTCC
AGGGATTTTTTAAAGTTAAATTTGGGAAATTAACAGCAGTTCACTTTATTGTGAGTCTTTGCCACATTTGACTGAATTTGAGCTGT
20 CATTTGTACATTTAAAGCAGCTGTTTTGGGCTCTGTGAGAGTACATGTATTATATACAAGCACAACAGGGCTTGCACTAAAGAA
TTGTCATTGTAATAACACTACTTTGGTAGCCTAACTTCATATATGTTATTCTTAATTGCACAAAAAGTCAATTAATTTGTACACCTG
GGGTTTTGAATGTTTGTCTTAAAGTGTGGCTATTTCTATGTTTATAAAACAAAACAAAATTTCCAAAAACAATGAAGGAAACC
AAAAATAATATTTCTGCATTTTCG
- 467
- 25 MADELEALRRQRLAELQAKHGDPGDAAQQEAKHREAEMRNSILAQVLDQSARARLSNLAIVKPEKTKAVENYLIQMARYGQLS
EKVSEQGLIEILKKVSQOTEKTTTVKFNRRKVMDSDEDDDY
- 468
- CTGCTCCAGCGCTGACGCCGAGCCATGGCGGACGAGGAGCTTGAGGCGCTGAGGAGACAGAGGCTGGCCGAGCTGCAGGCCAAA
CACGGGGATCCTGGTGATGCGGCCCAACAGGAAGCAAAGCACAGGGAAGCAGAAATGAGAAACAGTATCTTAGCCCCAAGTTCTG
30 GATCAGTCGGCCCCGGGCCAGGTTAAGTAACCTTAGCACTTGTAAGCCTGAAAAAATAAGCAGTAGAGAAATACCTTATACAG
ATGGCAAGATATGGACAACAAAGTGAAGGATCAGAACAAAGGTTTAAATAGAAATCCTTAAAAAGTAAGCCAACAAACAGAA
AAGACAACAACAGTGAAATTAACAGAGAAAAGTAATGGACTCTGATGAAGATGACGATTATTGAACTACAAGTGCTCACAGA
CTAGAACTTAACGGAACAAGTCTAGGACAGAAAGTTAAGATCTGATTATTTACTTTTATTGTCTATATGCTTTTAAAAAAA
TAAACTTGTATTGCAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
- 35 469
- MQVSTAALAVLLCTMALCNQVLSAPLAADTPTACCFSYTSRQIPQNFADYFETSSQCSKPSVIFLTRGRQVCADPSEEWQK
YVSDLELSA
- 470
- CGCAAAGAGTAGTCACTCCCTTCTTGGCTCTGCTGACACTCGAGCCACATTCATCACCTGCTCCCAATCATGCAGGTCTCCA
40 CTGCTGCCCTTGCCGTCCTCTCTGCACCATGGCTCTCTGCAACCAAGGTCTCTCTGCACCACGTGAGTCCATGTTGTTGTTGT
GGGTATCACCCTCTCTGGCCATGGTTAGACCAATCAGTCTTTTTCGCGGCTGAGAGCCCCGAAGAGAAAAGAAGGAAGTT
CTTAAAGCGCTGCCAAACACCTTGGTCTTTTCTTCAACATTTTATTTTATCTCTAGAAGGGGTCTTAGCCCTCCTAGTCTC
CAGTTGCTGCTGACACGCCGACCGCTGCTGCTTCAGCTACACCTCCCGACAGATTCCACAGAAATTCATAGCTGACTACTTTG
AGACGAGCAGCCAGTGCTCCAAGCCCAGTGTCATCTTCCTAACCAAGAGAGGCCGCGAGGTCTGTGCTGACCCCAAGTGAGGAGT
45 GGGTCCAGAAATACGTCACTGACCTGGAGCTGAGTGCTGAGGGGTCCAGAGCTTCGAGGCCACGACCTCAGTGGGCCAG
TGGGGAGGAGCAGGAGCTGAGCCTTGGGAACATGCGTGTGACCTCTACAGCTACCTCTCTATGAGCTGGTTATTGCCAAACA
GCCACACTGTGGGACTCTTCTTAACCTTAAATTTTAAATTTATTTTATACTATTTAGTTTTTATAATTTATTTTGTGATTTACAGTG
TGTTTGTGATTGTTTGTCTCTGAGAGTTCCCCCTGTCCCTCCACCTTCCCTCACAGTGTGTGCTGGTGACAACCGAGTGCTGCT

ATCGGCCGTGTAGGCAGTGATGGCACCAAGCCACCAGACTGACAAATGTGTGATCAGATGCTTTTGTTCAGGGCTGTGATCG
GCCTGGGGAAATAATAAGATGTTCTTTTAAACGGT

471

5 MELRPWLLWVVAATGTLVLLAADAQGGQKVFNTWAVRI PGGPAVANSVARKHGFNLGQIFGDYYHFWHRGVTKRSLSPHRPRH
SRLQREPQVQWLEQQVAKRRRTKRDVYQEPD PKFPQWYLSGVTRDLNVKAAWAQGYTG HGIVVSILDDGIEKNHPDLAGNYD
PGASFDVNDQDPDPQPRYTQMNDNRHGT RCAGEVA AVANNVCGVG VAYNARIGGVRLDGEVTD AVEARS LGLNPNHIH IYSA
SWGPEDDGKTVDGPARLAEEAFFRGVSQGRGGLSIFVWASNGNGREHDSNCNDGYTNSIYTL SISSATQFGNV PWYSEAC SST
LATYSSGNQNEKQIVTTDLRQKCTESH TGTSSASAPLAAGIIALTLEANKNL TWDRMQHLVVQTSKPAHLNANDWATNGVGRKV
10 SHSYGYGLLDAGAMVALAQNWTTVAPQRKCIIDILTEPKDIGRLEVRKTVTACLGEPNHITRLEHAQARLTLSYNRRGDLAIH
LVSPMGT RSTLLAARPHDYSADGFNDWAFMTTHSWDEDP SGEWVLEI ENTSEANNYGT LTKFTLVLYGTAP EGLFPVPPESSGCK
TLTSSQACV VCEEGFSLHQKSCVQHC PPGFAPQVLDTHYSTENDVETIRASVCA PCHASCATCQGPALTDCLSCPSHASLDPVE
QTCRSQS SRSRESPPQQQPPRLPPEVEAGORLRAGLLPSHLPEVVAGLSCAFI VLVFVTVFLVLQLRSGFSFRGVKVYTMDRGL
ISYKGLPPEAWQECPSDSEDEGRGERTAFIKDQSAL

472

15 GCGGGGAAGCAGCAGCGGCCAGGATGAATCCAGGTGCTCTGGAGCTGGATGGTGAAGGTCGGCACTCTTCACCTCCCGAGCC
CTGCCGCTCTCGGCCCCATGCCCCACCAGTCAGCCCCGGGCCACAGGCAGTGAGCAGGCACCTGGGAGCCGAGGCCCTATGAC
CAGGCCAAGGAGACGGGCGCTCCAGGGTCCAGCCACCTGTCCCCCATGGAGCTGAGGCCCTGGTTGCTATGGGTGGTAGCA
GCAACAGGAACCTTGGTCTGCTAGCAGCTGATGCTCAGGGCCAGAAGGTCTTCAACCAACAGTGGGCTGTGCGCATCCCTGGA
GGCCAGCGGTGGCCAACAGTGTGGCAGCAAGCATGGGTTCCTCAACCTGGGCCAGATCTTCGGGGACTATTACCACCTCTCTGG
20 CATCGAGGAGTGACGAAGCGGTCCCTGTGCGCTCACCGCCCGCGGCACAGCCGGCTGCAGAGGGAGCCCTCAAGTACAGTGGCTG
GAACAGCAGGTGGCAAAGCGACGGACTAAACGGGACGTGTACCAGGAGCCACAGACCCCAAGTTCTTCAGCAGTGGTACCTG
TCTGGTGTCACTCAGCGGGACCTGAATGTGAAGCGCGCTGGGCGCAGGGCTACACAGGGCACGGCATTTGTGGTCTCCATTCTG
GACGATGGCATCGAGAAGAACCACCGGACTTGGCAGGCAATTATGATCCTGGGGCCAGTTTGTATGTCAATGACCAGGACCT
GACCCCCAGCCTCGGTACACACAGATGAATGACAACAGGCACGGCACACGGTGTGCGGGGAAGTGGCTGCGGTGGCCAACAAC
25 GGTGTCTGTGGTGTAGGTGTGGCTTACAACGCCCGCATTTGGAGGGGTGCGCATGCTGGATGGCGAGGTGACAGATGCACTGGAG
GCACGCTCGCTGGGCTGAACCCCAACCACATCCACATCTACAGTGCAGCTGGGGCCCCGAGGATGACGGCAAGACAGTGGAT
GGGCCAGCCCGCTCGCCGAGGAGGCTTCTTCCGTGGGGTTAGCCAGGGCCGAGGGGGCTGGGCTCCATCTTTGTCTGGGCC
TCGGGGAACGGGGCCGGGAACATGACAGCTGCAACTGCGCAGGCTACACCAACAGTATCTACAGCTGTCCATCAGCAGCGCC
ACGCAGTTTGGCAACGTGCGGTGTACAGCGAGGCTGCTGCTCCACACTGGCCACGACCTACAGCAGTGGCAACCAGAATGAG
30 AAGCAGATCGTAGCAGTGA CTGCGGCAGAAGTGCACGGAGTCTCACAGGGCACCTCAGCCTCTGCCCTTAGCAGCCGGC
ATCATTGCTCTCACCTGGAGGCCAATAAGAACCTCACATGGCGGGACATGCAACACCTGGTGGTACAGACCTCGAAGCCAGCC
CACCTCAATGCCAACGACTGGGCCACCAATGGTGTGGGCGGAAAGTGAGCCACTCATATGGCTACGGGCTTTTGGACGAGGC
GCCATGGTGGCCCTGGCCAGAATGGACACAGTGGCCCCCAGCGGAAGTGATCATCGACATCTCACCAGGCCCAAGAC
ATCGGGAACGGCTCGAGGTGCGGAAGACCGTGACCGCTGCTGGGCGAGCCCAACCACATCACTCGGCTGGAGCAGGCTCAG
35 GCGCGGCTCACCTGTCTTATAATCGCCGTGGCGACCTGGCCATCCACCTGGTCAGCCCATGGGCACCGCTCCACCTGCTG
GCAGCCAGGCCACATGACTACTCCGCAGATGGGTPTAATGACTGGGCTTCATGACAACCTATTCTGGGATGAGGATCCCTCT
GGCGAGTGGGTCTAGAGATTGAAAACACCAGCGAAGCCAACAATATGGGACGCTGACCAAGTTCACCTCGTACTCTATGGC
ACCGCCCCTGAGGGGTGCCCCGTACCTCCAGAAAGCAGTGGCTGCAAGACCTCACGTCCAGTCAGGCTGTGTGGTGTGCGAG
GAAGGCTTCTCCCTGCACCAAGAGCTGTGTCCAGCACTGCCCTCCAGGCTTCGCCCCCAAGTCTCGATACGCACTATAGC
40 ACCGAGAATGACGTGGAGACCATCCGGCCAGCGTCTGCGCCCCCTGCCACGCTCATGTGCCACATGCCAGGGGGCGGCCCTG
ACAGATGCCTCAGCTGCCCCAGCCACGCTCTCTGGACCTGTGGAGCAGACTTGTCTCCGGCAAAGCCAGAGCAGCCGAGAG
TCCCCGCCACAGCAGCAGCCACCTCGGCTGCCCCCGAGGTGGAGGCGGGGCAACGGCTGCGGGCAGGGCTGCTGCCCTCACAC
CTGCTGAGGTGGTGGCGGCTCAGCTGCGCTTCATCGTGTGTCTTCTGCTACTGTCTTCTGCTGCTGAGCTGCGCTCT
GGCTTTAGTTTTCGGGGGTGAAGGTGTACACCATGAGACCGTGGCTTCCTTACAAAGGGGCTGCCCTTGAAGCTGGCAG
45 GAGGAGTGGCGCTGACTCAGAAGAGGACGAGGGCGGGGCGAGAGGACCGCTTTATCAAAGACCAAGACGCGCTCTGATGA
GCCCCACTGCCACCCCCCTCAAGCCAATCCCTCTTGGGCACTTTTAAATTCACCAAGTATTTTATCTTGGGACTGGGTT
TGGACCCACAGTGGGAGGCAAGAGGGGTGGAGACTGTTTCCATCTTACCTCGGGCCACCTGGCCACCTGAGGTGGGCCAG
GACCAGCTGGGGCGTGGGAGGGCGGTACCCACCTCAGCACCTTCCATGTGGAGAAAGGAGTGAACCTTTAGGCGAGCT
TGCCCCGGCCCCGGCCCCAGCCAGAGTTCTGCGGAGTGAAGAGGGGCGAGCCCTTGCTTGTGGGATTCCTGACCCAGGCCGCA
50 GCTCTTGCCCTTCCCTGTCCCTCTAAAGCAATAATGGTCCCATCCAGGCAGTCGGGGCTGGCTAGGAGATATCTGAGGGAGG

- AGGCCACCTCTCCAAGGGCTTCTGCACCCTCCACCCTGTCCCCAGCTCTGGTGAGTCTTGGCGGCAGCAGCCATCATAGGAAG
GGACCAAGGCAAGGCAGGTGCCTCCAGGTGTGCACGTGGCATGTGGCCTGTGGCCTGTGTCCATGACCCACCCTGTGCTCCG
TGCCTCCACCACCACTGGCCACCAGGCTGGCGCAGCCAAGGCCGAAGCTCTGGCTGAACCTGTGTGGTGTCTGACCAACCT
CCCCCTCTTGTGACCCGCCCTCTCCCGTCAGGGCCCAAGTCCCTGTTTCTGAGCCCGGGCTGCCCTGGGCTGTGGCACTCACAG
5 ACCTGGAGCCCCTGGGTGGGTGGTGGGGAGGGGCGCTGGCCAGCCGGCCTCTCTGGCCTCCACCCGATGTGCTTTCCCTTG
TGGGGATCTCAGGGGCTGTTTGGAGATATATTTTCACTTTGTGATTATTTTCACTTTAGATGTGATGATTTGTTTTGTATTTT
TAATGGGGTAGCAGCTGGACTACCCACGTTCTCACACCACCGTCCGCCCTGCTCCTCCCTGGCTGCCCTGGCCCTGAGGTGT
GGGGGCTGCAGCATGTTGCTGAGGAGTGAGGAATAGTTGAGCCCCAAGTCTGAAGAGCGGGCCAGCCAGGCGGGCTCAAGGA
AAGGGGCTCCAGTGGGAGGGGAGGCTGACATCTGTGTTTCAACTGGGGCTCGCCATGCCGGGGGTTTCATAGGTCACCTGGCTC
10 TCCAAGTGCCAGAGGTGGGAGGTGGTGGCACTGAGCCCCCAACACTGTGCCCTGGTGGAGAAAGCACTGACCTGTCATGCC
CCCTCAAACCTCCTCTTCTGACGTGCCCTTTTGCACCCCTCCCATTAGGACAATCAGTCCCTCCCATCTGGGAGTCCCTTTT
CTTTTCTACCTTAGCCATTCTGTGACCCAGCCATCTGCCAGGGGTGCCCCCTCTCTCCATCCCCCTGCCCTCGTGCCAG
CCCGGCTGGTTTTGTAAGATACTGGGTGGTGCACAGTGATTTTTTCTTGTAAATTTAAACAGGCCAGCATTGCTGGTTCTAT
TTAATGGACATGAGATAATGTTAGAGGTTTAAAGTGATTAAACGTGCAGACTATGCAAACCAG
- 15 473
MAMESTATAAADVVSADKIEDVPAPSTSADKVESLDVDSEAKLLGLGQKHLVMGDI PAAVNAFQEAASLLGKKYGETANEC
GEAFFFYGKSLLELARMENGLVNALEGVHVEEBEKEKTEDESLENNDNIDEEAREELREQVYDAMGEKEAKKTEDKSLAKP
ETDKEQDSEMEKGGREDMDISKSABEPQEKVDLTLDWLTETSEEAKGGAPEGPNEAEVTSKGPEQEVPADEEEKSVSGTDVQE
ECREKGGQEKQGEVIVSIEEKPEVSEEQPVVLEKQGTAVEVEAESLDPTVKPVDVGGDEPEEKVVTSENEAGKAVLEQLVGQ
20 EVPPAESPEVQTEAAEASAVEAGSEVSEKPGQEPVLPKDGAVNGPSVVGDTPIEPQTSIERLTETKDGSGLEEKVRAKLVP
SQEETKLSVEESEAAGDGVDTKVAQGATEKSPEDKVQIAANEETQEREEQMKEGEETEGSEEDDKENDKTEEMPNDVLENKSL
QENEEEEIGNLELAWDMLDLAKIIFKRQETKEAQLYAAQHLKLGEVSVESENYVQAVEEFQSCNLNQEQYLEAHRLLAETHY
QLGLAYGINSQYDEAVAQFSKSIIEVENRMAVLNEQVKEAESSEYKKEIEELKELLPEIREKIEDAKESQSRSGNVAELALKAT
LVESSTSGFTPGGGSSVSMIASRKPTDGASSNCVTDISHLVRRKKRPEEESPRKDDAKKAKQEPEVNGSGDAVPSGNEVSE
25 NMEEEAENQLKRGAAVEGTLEAGATVESTAC
474
GCCTGAGTGAGTCTCTGGCGTCCCAAATTGCCGTGTTTTCTCGCAGGCTCTATTCCGTTGCTGGTTCGCCACCTCAGGGGAAC
GATGGCCATGGAGTCCACAGCCACTGCCCGCGTCCGCCGGACGTGGTTTTCTGCCGACAAAATTGAAGATGTCCCTGCTCCTTC
TACATCTGCAGATAAAGTGGAGAGTCTGGATGTGGATAGTGAAGCTAAGAAAATATTGGGTTTAGGACAGAAACATCTGGTGAT
30 GGGGATATTCCAGCAGCTGTCAATGCATTCAGGAAGCAGCTAGTCTTTTAGGTAAGAATATGGAGAGACAGCTAATGAGTG
TGGAGAAGCCTTCTTTTCTATGGGAAATCACTTCTGGAGTTGGCAAGAATGGAGAATGGTGTGTTGGGAAACGCCCTTGAAGG
TGTGCATGTGGAAGAGGAAGAAGGAGAAAAACAGAAGATGAATCTCTGGTAGAAAATAATGATAACATAGATGAGGAAGCAAG
GGAAGAGTTGAGAGAACAGGTTTATGACGCCATGGGAGAAAAAGAAGAAGCCAAAAAACAGAAGACAAGTCTTTGGCAAAGCC
TGAAACTGATAAAGAACAGGACAGTGAATGGAGAAGGGTGAAGAGAAGATATGGATATAAGTAAATCTGCAGAGGAGCCACA
35 GGAAAAAGTTGACTTGACTCTAGATTGGTTAACTGAAACCTCTGAAGAGGCAAAAGGAGGACAGCACCAGAAGGACCGAATGA
AGCTGAGGTCACCTTCTGGGAAGCCAGAACAGGAAGTACCAGATGCTGAGGAAGAAAAATCAGTTTCTGGAAGTGTGTCGAAGA
AGAGTGCAGAGAAAAAGGAGGTCAGGAGAAGCAGGAGAGGTAATTGTGAGCATAGAGGAGAAGCCAAAAGAAGTTTCAGAAGA
GCAGCCTGTGGTGAATCTAGAAAAGCAGGGCACTGCAGTGGAGGTAGAAGCAGAGTCTTTAGACCCGACAGTCAAGCCAGTGGA
TGTGGGTGGGACGAGCCAGAGGAGAAGGTAGTTACCTCTGAAAACGAGGCAGGAAAGGCGGTCTTTGAACAACTGGTAGGTCA
40 AGAAGTACCACCTGCTGAAGAGTCAACAGAGGTGCAACAGAGGCTGCAGAGGCCCTCAGCTGTAGAGGCTGGATCAGAAGTCTC
TGAAAAGCCTGGGCAGGAGGCTCAGTTCTCCCTAAGGATGGTGCAGTCAATGGACCGTCAGTTGTAGGAGATCAGACTCCTAT
TGAAACCACAGACTTCTATAGAAAGACTGACAGAAAACAAAGATGGCTCAGGACTAGAGGAGAAGGTCAGGGCAAAGCTGGTTCC
TAGTCAGGAGGAGACTAAGCTGTCTGTAGAAGAGTCTGAGGCAGCTGGAGATGGGTTGATACCAAGGTAGCCAGGGAGCTAC
TGAGAAATCAGCTGAAGACAAAATTGAGATAGCTGCTAATGAAGAGACACAAGAGAGAGAAGAATGAAAGAGGGTGAAGA
45 AACTGAAGGCTCGGAAGAGGATGATAAAGAAAATGATAAGACTGAAGAAAATGCCAAATGATTGACCTTTGAAAACAAAGTCTCT
TCAAGAAAATGAGGAGGAGGAGATTGGGAACCTAGAGCTTGCTGGGATATGCTGGATTTAGCAAAGATCATTTTTAAAGGCA
AGAAACAAAAGAAGCACAGCTTTATGCTGCCAGGCACATCTTAACTCGGAGAAGTTAGTGTGAATCTGAAAACATATGTGCA
AGCTGTGGAGGAGTTCCAGTCTGCTTAACCTGCAGGAACAGTACCTGGAAGCCACGACCGTCTGCTGCAGAGACCCACTA
CCAGCTGGGCTTGGCTTATGGGTACAATCTCAGTATGATGAGGCAGTGGCACAGTTCAGCAAATCTATTGAAGTCATTGAGAA
50 CAGAATGGCTGTACTAAACGAGCAGGTGAAGGAGGTGAAGGATCGTCTGAATACAAGAAAGAAATGAGGAACATAAGGAACT

GCTACCCGAAATTAGAGAGAAGATAGAAGATGCAAAGGAGTCTCAGCGTAGTGGGAATGTAGCTGAACTGGCTCTGAAAGCTAC
TCTGGTGGAGAGTTCTACTTCAGGTTTCACTCCTGGTGGAGGAGGCTCTTCAGTCTCCATGATTGCCAGTAGAAAGCCAAACAGA
CGGTGCTTCTCATCAAAATTGTGTGACTGATATTTCCACCTTTGTCAGAAAAGAGAGGAAACCAGAGGAAGAGAGTCCCCGGAA
AGATGATGCAAAGAAAGCCAAACAAGAGCCGGAGGTGAACGGAGGCAGTGGGGATGCTGTCCGAGTGGAAATGAAGTTTCGGA
5 AAACATGGAGGAGGAGGCTGAGAATCAGCTGAAACGGAGCAGCAGTGGAGGGGACACTGGAGGCTGGAGCTACAGTTGAAAG
CACTGCATGTTAAGAGGGGGCACAGCCTCCTCCCAAGGAAAGTGTTTTGTATATAATGTATTTTTCACTTTGGAGGATTC
TTTTTGTATAACTTCAATAAAGATTGTAAGCAAAAAAAAAA

475

MEHTGHYHLHAPLMTTVFSLSPGKANYTRLWANSTSSWDSVIQNKGTGRNQENINTNPITPEVDYKGNSTNMPETSHIVALTS
10 KSEQELYIPSVSNSPSTVQSIENTSKSHGEIFKKDVCAENNNMAMLICLIIIAVLFLICTPLFLSTVVLANKVSSLRRSKQV
GKRQPRSNGLFLASGLWPAESDTWKRTKQLTGPNLVMQSTGVLATATREKDEEGTEKLTNKQIG

476

AGTTGCAGTGGAAAGAAATGTGTCATCTGTGGTTTGGTTTTTAAAAGTGGAAACTAGCTGCACATATCCTTTTTTACTGCAGA
TTTACTTTAAGGCTCATATCTCCAAGTCTATTCGCTTTAAAAGAAGACAAGAAAAGAGTGGTTTATCAAAATCACGTTAT
15 AATCAGATTTTGACCAAGCATTTTGTAAAGATTGCCAAGTATGCCACGGACATGGAACACACAGGACATTACCTACATCTTGCC
TTTCTGATGACAACAGTTTTCCTTTGTCTCCTGGAACAAAAGCAAACATAACCCGCTCTGTGGGCTAACAGTACTTCTCCTGG
GATTCACTTATTTCAAAACAAGACAGGCAGAAACCAAAATGAAAACATTAAACACAAACCTATAACTCCTGAAGTAGATTATAAA
GGTAATTCTACAAACATGCCTGAAACATCTCACATCGTAGCTTTAACTTCTAAATCTGAACAGGAGCTTTATATACCTTCTGTC
GTCAGCAACAGTCTTCAACAGTACAGAGCATTGAAAACACAAGCAAAGTCAATGGTGAATTTTCAAAAAGGATGTCTGTGCG
20 GAAAACAACAACATGGCTATGCTAATTTGCTTAATTATAATTGCACTGCTTTTCTTATCTGTACCTTTCTATTCTATCA
ACTGTGGTTTGGCAAACAAGTCTCTCTCTCAGACGATCAAAACAAGTAGGCAAGCGTCAGCCTAGAAGCAATGGCGATTTT
CTGGCAAGCGGCTATGGCCCGCTGAATCAGACACTTGGAAAAGAACAACAGCTCACAGGACCAACCTAGTGATGCAATCT
ACTGGAGTGCTCACAGCTACAAGGAAAGAAAAGATGAAGAAGGAAGTGAAGTGGTATTTAATCCCAAAGTGTGTTCTGATTATC
AAAATGCAAAGTAGCAATGAGAAGCCTTATGGAGTAAAAATGAAGTCAAGTGGTATTTAATCCCAAAGTGTGTTCTGATTATC
25 TAAAATTTGACATGGTAGACCTTGCAATTTAGAAATCAAGCAGGTGAGACAGGGAGAAGTATGCCCTGCTTAATTTAACTGAT
GTACTTTTGTGTTGACACTGAATATTTTAAAAAGCAAATAATAAAATAACTAAGCATTTGAGGAAAAATTTAAGGATAAATTTGA
GGAACTGATTAATAGAGATAGCAAGGGATAATTAATAAATATTCCTATGTAGCAACAGTGGTTAGATGATCTTTGTCTGAA
TGTAATAAACTTTGAATAGTTTGTAGTGTCTTAAAGCCAAGTATATGCTTTAATCAATGGAAGTCAAAATTCCTAATGC
ATAGATAGAGAGCTAACTGTGTAATTTAATGGTATCTTCTTGTGATGTTGGCAGAAATCCACACCAGCTTATCAACCAAC
30 ACAGCTAATTTTGAATAGGTCTTTATCTTTCCATATGGCACAGTAAAGAAAGTGTTTTCTACTATTAAATTTAAAA
CCTTTACTTTTGTATAATAAATTTAACTCAGAATAAACCTGTGACCACGT

477

MAENGDNKMAALEAKICHQIEYYFGDFNLPRDKFLKEQIKLDEGWVPLEIMIKFNRLNRLTTDFNVIVEALSKSKAELMEISE
DKTKIRRSPLPEVTDDEYKNDVKNRSVYIKGPPTDALTDDIKEWLEDKGQVLNIQMRRTLHKAFFKSIFVVFDSIESAKKPV
35 ETPGQKYKETDLLILFKDDYFAKNEERKQNKVEAKLRAKQEQAKQKLEEDAEMKSLEEKIGCLLKPSGDLDDQTCREDLHIL
FSNHGEIKWIDFVRGAKEGIIILFKEKAKEALGKAKDANNGNLQLRNKEVTWEVLEGEVEKEALKKIIEDQESLNKWKSKGRRF
KGKKGKNAAPGSGKGKVQFQKKTKFASDDEHDEHENGATGPVKRAREETDKEEPASKQKKTENGAGDQ

478

GGAGTCGTTGTTGTGCTGTTTGTGAGCCTGTGCGGCGGCTTCTGTGGGCCGGAACCTTAAAGATAGCCGAATGGCTGAAAT
40 GGTGATAATGAAAAGATGGCTGCCCTGGAGGCCAAAATCTGTCATCAAAATTGAGTATTATTTTGGCGACTTCAATTTGCCACGG
GACAAGTTTCTAAAGGAACAGATAAACTGGATGAAGCTGGGTACCTTTGGAGATAATGATAAAATTAACAGGTTGAACCGT
CTAACAACAGACTTTAATGTAATTGTGGAAGCATTGAGCAATCCAAGGCAGAACTCATGGAATCAGTGAAGATAAACTAAA
ATCAGAAGTCTCCAAGCAAACCCCTACCTGAAGTACTGATGAGTATAAAATGATGTAAAAAACAGATCTGTTTATATTAAA
GGCTTCCCACTGATGCAACTTTGATGACATAAAGAATGGTTAGAAGATAAAGTCAAGTACTAATATTAGATGAGAAGA
45 ACATTGCATAAAGCATTAAAGGATCAATTTTGTGTTGTTGATAGCATTGAATCTGCTAAGAAATTTGTAGAGACCCCTGGC
CAGAAGTACAAAGAAACAGACCTGCTAATACTTTTCAAGGACGATTACTTTGCCAAAAAAATGAAGAAAGAAAACAAAATAAA
GTGGAAGCTAAATTAAGAGCTAAACAGGAGCAAGAAGCAAAACAAAAGTTAGAAGAAGATGCTGAAATGAAATCTCTAGAAGAA
AAGATTGGATGCTTGTGAAATTTTCGGGTGATTTAGATGATCAGACCTGTAGAGAAGATTTACACATACTTTTCTCAAATCAT

[illegible]

- CGATAATGTCTCTCAAGATTTCAAAGTCATATGAGATTGGGATATTTTTGTACAGGTTGTGTTTGTATTATGTCAGTTTTTAA
TAAACATAAATGTGGGACAGAGCTGTCTATTTAGTATATCAAAGTTTAGTAGTTTCTCCACATTCACGAAATTACCACAGTG
AGAGCTAAGCATTTCTACTGGGCAGTTTCATTTTGTAGTTGATCAGGTTTAAAGTTTTTGAACATAAAATTTCTTTTCTTTTT
ATGATGAATAAGGTTAAAATAAAAGCCTTAGACAAAATAAATTTGGCAGAGTTAATTGAGCAAAGGACAATTCACAAATCAGG
5 TAGCCCCGTAACCATAATAGGCTCAGAGGCTTCAGCCAGCTGCATAGTTGAAGATTATGGACAGAAGGAAAGTGATGTATGG
AAAATGGAAGTGAGATACAGCAACAGCCGGATTAGTTACAGTTTCAGCGTTTGCCCTTATTTGAATATGGTTTGAACAGTTTCGCTG
TCTTTGGTTGGCTGAACTTAGTGATTGCCACAAGAGTAGGGTACCGTCTGTTTACACGTCCAGTTAGGCTACAGTTCTATGTA
CTGAGAAACCTTTAAGCTGAACCTGAGATATGTAAAGAGACTTTAGGCTAACTTAAACAATATATATAGGAATATATCCCTTCT
ACTTCACATGCACTGAATATGCATTTTATTGCTTTACTCTTCATTCTGTGGCACCTACCCACAGGGGAAGTAAGAAGTTGTTT
10 TGGTATTTCCGAAACTAAAGTCCTTATGGGATGGGGTCTAGAATTGATTCTCCTTTCCCTGAGTTTACTCCACGGAGTCTTAGG
TACCTGGTAAAAAGTTGCTTCTTAAATTAAGGGTCATTGCTTTGTTGCTAGCTGCTAATGCTTACCTTTGTTTCTTTGCTT
TTTAATCAGTTCTTAATAGGATATAGTTTATGTTTCCAAGTTATAACTTGGAGTTAATGGTCACTAGATTATCAGTTATGAG
CAGTGTTAAAAATCTCCTATTAATGTGTAATGTACCTGTGACGTCCCTCTTATTAAGGGTTCTTTGAGAATAAAAGAGAAAAG
ACCTACTTTATTTGACAGCAAAAAAAGGAATTC
- 15 485
MARAALSAAPSNPRLLRVALLLLLLVAAGRRAAGASVATELRQCQLQTLQGIHPKNIQSVNVKSPGPHCAQTEVIATLKNRKA
CLNPASPIVKKIIKMLNSDKSN
486
CACAGAGCCCGGCCGAGGCACCTCCTCGCCAGCTCTTCCGCTCCTCTCACAGCCGCCAGACCCGCTGCTGAGCCCCATGGC
20 CCGCGCTGCTCTCTCCGCGCCCCCAGCAATCCCGGCTCCTGCGAGTGGCACTGCTGCTCCTGCTCCTGGTAGCCGCTGGCCG
GCGCGCAGCAGGAGCGTCCGTGGCCACTGAACCTGCGCTGCCAGTGTCTGCAGACCTGCAGGGAATTCAACCCAAAGAACATCCA
AAGTGTGAACGTGAAGTCCCCCGACCCACTGCGCCCAACCGAAGTCATAGCCACACTCAAGAATGGCGGAAAGCTTGCCT
CAATCCTGCATCCCCATAGTTAAGAAAATCATCGAAAAGATGCTGAACAGTGACAAATCCAAGTACCAGAAGGGAGGAGGAA
GCTCACTGGTGGCTGTTTCTGAAGGAGGCCCTGCCCTTATAGGAACAGAAGAGGAAAGAGAGACACAGCTGCAGAGGCCACCTG
25 GATTGTGCCTAATGTGTTTGAGCATCGCTTAGGAGAAGTCTTCTATTTATTTATTTATTCATTAGTTTTGAAGATTCTATGTTA
ATATTTTAGGTGAAAAATAATTAAGGGTATGATTAACTTACCTGCACACTGTCTATATATTCATTCTTTTGAAGTGTCAA
CCCCAAGTTAGTTCAATCTGGATTATATTTAATTTGAAGGTAGAATGTTTTCAAAATGTCTCCAGTCAATTATGTTAAATTTTC
TGAGGAGCTGCAACATGCCAGCCACTGTGATAGAGGCTGGCGGATCCAAGCAATGGCCAATGAGATCATTTGTAAGGCAGGG
GAATGTATGTGCACATCTGTTTGTAACTGTTTATAGTGAATGTGAGTTGTTATTTATTTGAAATGATTTCACAGTGTGTGGTCAA
30 CATTTCTCATGTTGAAACTTTAAGAACTAAAATGTTCTAAATATCCCTTGGACATTTTATGCTTTCTTGTGAAGGCATACTGCC
TTGTTTAAATGGTAGTTTACAGTGTCTTGGCTTAGAACAAAGGGGCTTAATTATTGATGTTTTCATAGAGAATATAAAAAATA
AGCACTTATAG
487
MIFPWKCQSTQRDLWNIIFKLWGWMLCCDFLAHGTDCWYHYSEKPMNWQRARRFCRDNYDLVAIQNKAEIEYLEKTLPFSSR
35 SYWIGIRKIGGIWTVWGTNKSLEEAENWGDGEPNNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKAALCYTASCQPWSCSGH
GECVEIINNYTCNCDVGYGPPQCQFVIQCEPLEAPELGMTDCTHPLGNFSFSSQCAFSCSEGTNLGTIEETTCCGPFGNWSSPEP
TCQVIQCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTELIGKKKTICESSGIWSNPSPICQKLDKSFSMIKEGDYNPLFI
PVAVMVTAFSGLAFIIWLARRLKKGKSKRSMNDPY
488
40 CTCCCTTTGGGCAAGGACCTGAGACCTTGTGCTAAGTCAAGAGGCTCAATGGGCTGCAGAAGAACTAGAGAAGGACCAAGCAA
AGCCATGATATTTCCATGGAAATGTCAGAGCACCCAGAGGACTTATGGAACATCTTCAAGTTGTGGGGGTGGACAATGCTCTG
TTGTGATTTCCTGGCACATCATGGAACCGACTGCTGGACTTACCATTATTTCTGAAAAACCCATGAAGTGGCAAAGGGCTAGAAG
ATTCTGCCGAGACAATTACACAGATTTAGTTGCCATACAAAACAAGGCGGAAATGAGTATCTGGAGAAGACTTGCCTTTTACAG
TCGTTCTTACTACTGGATAGGAATCCGGAAGATAGGAGGAATATGGACGTGGGTGGGAACCAACAAATCTTACTGAAGAAGC
45 AGAGAAGTGGGAGATGGTGAGCCCAACAACAAGAACAAGGAGGACTGCGTGGAGATCTATATCAAGAGAAACAAGATGC
AGGCAAATGGAACGATGACGCTGCCACAACTAAAGGCAGCCCTCTGTTACACAGCTTCTTGGCAGCCCTGGTCAAGCAGTGG
CCATGGAGAATGTGTAGAAATCATCAATAATTACACCTGCAACTGTGATGTGGGTACTATGGGCCCCAGTGTGAGTTTGTGAT
TCAGTGTGAGCCTTTGGAGGCCCCAGAGCTGGGTACCATGGACTGTACTACCCCTTTGGGAAACTTCAGCTTCAGCTCACAGTG

TGCCTTCAGCTGCTCTGAAGGAACAACTTAACTGGGATTGAAGAAACCACCTGTGGACCATTTGGAACTGGTCATCTCCAGA
ACCAACCTGTCAAGTGATTCAGTGTGAGCCTCTATCAGCACCAGATTTGGGGATCATGAACTGTAGCCATCCCCTGGCCAGCTT
CAGCTTTACCTCTGCATGTACCTTCATCTGCTCAGAAGGAACAGTAAATTGGGAAGAAGAAAACCATTTGTGAATCATCTGG
AATCTGGTCAAATCCTAGTCCAATATGTCAAAAATTGGACAAAAGTTTCTCAATGATTAAGGAGGGTGATTATAACCCCTCTT
5 CATTCAGTGGCAGTCATGGTTACTGCATTCTCTGGGTTGGCATTATCATTTGGCTGGCAAGGAGATTAAAAAAGGCAAGAA
ATCCAAGAGAAGTATGAATGACCCATATTAATCGCCCTTGGTGAAAGAAAATTCTTGGAACTATAAAAAATCATGAGATCCTTT
AAATCCTTCCATGAAACGTTTGTGTGGTGGCACCTCCTACGTCAAACATGAAGTGTGTTTCCCTTCAGTGCATCTGGGAAGATT
TCTACCTGACCAACAGTTCCCTTCAGCTTCCATTTGCCCCCTCATTTATCCCTCAACCCCAAGCCACAGGTGTTTATACAGCTC
AGCTTTTGTCTTTTCTGAGGAGAAACAAATAAGACCATAAAGGGAAGGATTCATGTGGAATATAAAGATGGCTGACTTTGCT
10 CTTTCTTGACTCTTGTTTTCAGTTTCAATTTCAGTGTGACTTGATGACAGACACTTCTAAATGAAGTCAAATTTGATACATA
TGTGAATATGGACTCAGTTTCTTGCAGATCAAATTTACGTCGTCTTCTGTATACTGTGGAGGTACACTCTTATAGAAAGTTC
AAAAAGTCTACGCTCTCCTTTCTTTCTAACTCCAGTGAAGTAATGGGGTCCTGCTCAAGTTGAAAGAGTCTTATTTGCACTGTA
GCCTCGCCGCTGTGAATTGGACCATCCTATTTAACTGGCTTCAGCCTCCCCACCTTCTTCAGCCACCTCTCTTTTTTCAGTTGG
CTGACTTCCACACCTAGCATCTCATGAGTGCCAAAGCAAAAGGAGAGAAGAGAGAAATAGCCTGCGCTGTTTTTTAGTTTGGGGG
15 TTTTGTCTGTTTCTTTTATGAGACCCATTCCTATTTCTTATAGTCAATGTTTCTTTTATCACGATATTATTAGTAAGAAAACAT
CACTGAAATGCTAGCTGCAAGTGACATCTCTTTGATGTCAATGGAAGAGTTAAAACAGGTGGAGAAATTCCTTGATTCACAAT
GAAATGCTCTCTTTTCCCTGCCCCCAGACCTTTTATCCACTTACCTAGATTCTACATATTCCTTAAATTTTCATCTCAGGCCTC
CCTCAACCCCAACACTTCTTTTATAACTAGTCCCTTACTAATCCAACCCATGATGAGCTCCTCTTCTGGCTTCTTACTGAAAG
GTTACCCGTGAACATGCAATTTTGCATTTGAATAAAGCCTGCTTTTTTAAGTGTTAA

CLAIMS

1. A method for the identification of a gene that is implicated in a specific disease or physiological condition, said method comprising the steps of:
 - 5 a) comparing:
 - i) the transcriptome or proteome of a first specialised cell type that is implicated in the disease or condition under first and second experimental conditions; with
 - ii) the transcriptome or proteome of a second specialised cell type under said first and said second experimental conditions; and
 - 10 b) identifying as a gene implicated in the disease or physiological condition, a gene that is differentially regulated in the two specialised cell types under the first and second experimental conditions.
2. A method according to claim 1, wherein said specialised cell types are selected from the group consisting of cardiomyocytes, endothelial cells, sensory neurons, motor neurons, CNS neurons,
15 astrocytes, glial cells, schwann cells, mast cells, eosinophils, smooth muscle cells, skeletal muscle cells, pericytes, lymphocytes, tumor cells, monocytes, macrophages, foamy macrophages, granulocytes, synovial cells / synovial fibroblasts and epithelial cells.
3. A method according to claim 1 or claim 2, wherein said first and second experimental conditions differ in respect of the cellular microenvironment, or in respect of exposure to hormones, growth factors,
20 cytokines, chemokines, inflammatory agents, toxins, metabolites, pH, pharmaceutical agents, hypoxia, anoxia, ischemia, imbalance of any plasma-borne nutrient, osmotic stress, temperature, mechanical stress, irradiation, cell-extracellular matrix interactions, cell-cell interactions, accumulations of foreign or pathological extracellular components, intracellular and extracellular pathogens, or a genetic perturbation.
- 25 4. A method according to any one of the preceding claims, wherein the first experimental conditions and second experimental conditions differ in that under the second experimental conditions, the cells are exposed to a physiological stimulus.
5. A method according to claim 4, wherein the physiological stimulus is a physiological, mechanical, temperature, chemical, toxic or pharmaceutical stress.
- 30 6. A method according to claim 5, wherein said physiological stress is hypoxia.

7. A method according to any one of the preceding claims, wherein said first and second experimental conditions are different genetic conditions.
8. A method according to claim 7, wherein said second experimental conditions differ from said first experimental conditions in that the expression of a genetic element is expressed at a different level in
5 said second experimental conditions relative to the level of expression of the genetic element in said first experimental conditions.
9. A method according to claim 8, wherein said genetic element is heterologous to the specialized cell type.
10. A method according to any one of the preceding claims, wherein the transcriptomes of the specialized
10 cell types are compared by a technique involving hybridization to a nucleic acid array, subtractive mRNA hybridisation, the serial analysis of gene expression (SAGE); the selective amplification via biotin- and restriction-mediated enrichment (SABRE); differential display; representational difference analysis (RDA); differential screening of cDNA libraries; Northern blotting; an RNase protection assay; an S1-nuclease protection assays; RT-PCR; real time RT-PCR (Taq-man); EST sequencing;
15 massively parallel signature sequencing (MPSS); or sequencing by hybridisation (SBH).
11. A method according to claim 10, wherein the transcriptomes are compared by hybridization to a nucleic acid array.
12. A substantially purified polypeptide, encoded by a gene implicated in a specific disease or physiological condition by a method according to any one of the preceding claims.
- 20 13. A substantially purified polypeptide, which polypeptide:
 - i) comprises the amino acid sequence as recited in any one of SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 63, 67, 69, 73, 75, 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129, 131, 133, 137, 139, 141, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 and 209;
 - 25 ii) has an amino acid sequence encoded by a nucleic acid sequence recited in any one of SEQ ID Nos: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164,
30 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200,

431

202, 204, 206, 208, 210, 212, 214 and 216, or has an amino acid sequence encoded by a gene identified from an EST recited in any one of these SEQ ID Nos;

- 5 iii) is a fragment of a polypeptide according to i) or ii), provided that said fragment retains a biological activity possessed by the full length polypeptide of i) or ii), or has an antigenic determinant in common with the polypeptide of i) or ii); or
- iv) is a functional equivalent of a polypeptide of i), ii) or (iii).

14. A polypeptide according to claim 13, wherein said biological activity is a hypoxia-regulated activity.

15. A polypeptide according to claim 14, wherein the expression of the polypeptide is hypoxia-induced.

16. A polypeptide according to claim 15, which polypeptide:

- 10 i) comprises the amino acid sequence as recited in any one of SEQ ID Nos.: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 63, 67, 69, 73, 75, 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129, 131, 133, 137, 139 and 141;
- ii) has an amino acid sequence encoded by a nucleic acid sequence recited in any one of SEQ ID Nos.: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 64, 66, 68, 70, 72, 74, 15 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142 and 144, or is encoded by a gene identified from an EST recited in any one of these SEQ ID Nos.;
- iii) is a fragment of a polypeptide according to i) or ii), provided that said fragment retains a biological activity possessed by the full length polypeptide of i) or ii), or has an antigenic 20 determinant in common with the polypeptide of i) or ii); or
- iv) is a functional equivalent of a polypeptide of i), ii) or (iii).

17. A polypeptide according to claim 14, wherein the expression of the polypeptide is hypoxia-repressed.

18. A polypeptide according to claim 17, which polypeptide:

- 25 i) comprises the amino acid sequence as recited in any one of SEQ ID Nos.: 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 and 209;
- ii) has an amino acid sequence encoded by a nucleic acid sequence recited in any one of SEQ ID Nos.: 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192,

194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216, or is encoded by a gene identified from an EST recited in any one of these SEQ ID Nos.;

iii) is a fragment of a polypeptide according to i) or ii), provided that said fragment retains a biological activity possessed by the full length polypeptide of i) or ii), or has an antigenic determinant in common with the polypeptide of i) or ii); or

iv) is a functional equivalent of a polypeptide of i), ii) or (iii).

19. A polypeptide which is a functional equivalent according to part iv) of any one of claims 13-18, is homologous to the amino acid sequence as recited in any one of SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 63, 67, 69, 73, 75, 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129, 131, 133, 137, 139, 141, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 and 209 or is homologous to the amino acid sequence encoded by a nucleic acid as recited in any one of SEQ ID Nos: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216, and has equivalent biological activity to that possessed by the full length polypeptide of i) or ii).

20. A fragment or functional equivalent according to any one of claims 13-19, which has greater than 50% sequence identity with the amino acid sequence as recited in any one of SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 63, 67, 69, 73, 75, 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129, 131, 133, 137, 139, 141, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 and 209 or with the amino acid sequence that is encoded by a nucleic acid as recited in any one of SEQ ID Nos: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216, or with fragments thereof, preferably greater than 60%, 70%, 80%, 90%, 95%, 98% or 99% sequence identity.

21. A fragment as recited in any one of claims 13-20 having an antigenic determinant in common with a polypeptide according to part i) of any one of claims 13-18, which consists of 7 or more (for example,

- 8, 10, 12, 14, 16, 18, 20 or more) amino acid residues from the amino acid sequence as recited in any one of SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 63, 67, 69, 73, 75, 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129, 131, 133, 137, 139, 141, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 and 209 or
- 5 the amino acid sequence encoded by a nucleic acid as recited in any one of SEQ ID Nos: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192,
- 10 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216.
22. A purified and isolated nucleic acid molecule that encodes a polypeptide according to any one of claims 13-21.
23. A purified nucleic acid molecule according to claim 22, which consists of the nucleic acid sequence as recited in any one of SEQ ID Nos.: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38,
- 15 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214 and 216, or is a redundant equivalent or fragment thereof.
- 20 24. A purified nucleic acid molecule which hybridizes under high stringency conditions with a nucleic acid molecule according to claim 22 or claim 23.
25. A vector comprising a nucleic acid molecule as recited in any one of claims 22-24.
26. A delivery vehicle comprising a nucleic acid according to any one of claims 22-24 or a vector according to claim 25.
- 25 27. A host cell transformed with a vector according to claim 25.
28. An antagonist ligand which binds specifically to a polypeptide according to any one of claims 13-21, preferably a ligand which inhibits the hypoxia-induced activity of said polypeptide.
29. An agonist ligand which binds specifically to a polypeptide according to any one of claims 13-21, preferably a ligand which augments or potentiates a hypoxia-induced activity of said polypeptide.
- 30 30. A ligand according to claim 28 or claim 29, which is an antibody.

31. A ligand according to claim 28 or claim 29, which is a peptide, a peptidomimetic, or a drug molecule, such as a small natural or synthetic organic molecule of up to 2000Da, preferably 800Da or less.
32. A polypeptide according to any one of claims 13-21, a nucleic acid molecule according to any one of claims 22-24, a vector according to claim 25 or a ligand according to claim 30 or 31, for use in therapy or diagnosis of disease.
33. A polypeptide, nucleic acid molecule, vector or ligand as recited in claim 32, wherein said disease is a hypoxia-regulated condition.
34. A polypeptide, nucleic acid molecule, vector or ligand as recited in claim 33, wherein said hypoxia-regulated condition is tumourigenesis, angiogenesis, apoptosis, inflammation, erythropoiesis, the biological response to hypoxia conditions (including processes such as glycolysis, gluconeogenesis, glucose transportation, catecholamine synthesis, iron transport or nitric oxide synthesis).
35. A substantially purified polypeptide, which polypeptide:
- i) comprises the amino acid sequence as recited in any one of SEQ ID Nos: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 63, 67, 69, 73, 75, 77, 85, 87, 89, 91, 93, 95, 99, 103, 113, 115, 119, 121, 129, 131, 133, 137, 139, 141, 145, 151, 153, 157, 159, 163, 169, 181, 187, 201, 205, 207 and 209 or SEQ ID Nos.: 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 239, 241, 243, 245, 247, 249, 251, 253, 255, 257, 259, 261, 263, 265, 267, 269, 271, 273, 275, 277, 279, 281, 283, 285, 287, 289, 291, 293, 295, 297, 299, 301, 303, 305, 307, 309, 311, 313, 315, 317, 319, 321, 323, 325, 327, 329, 331, 333, 335, 337, 339, 341, 343, 345, 347, 349, 351, 353, 355, 357, 359, 361, 363, 365, 367, 369, 371, 373, 375, 377, 379, 381, 383, 385, 387, 389, 391, 393, 395, 397, 399, 401, 403, 405, 407, 409, 411, 413, 415, 417, 419, 421, 423, 425, 427, 429, 431, 433, 435, 437, 439, 441, 443, 445, 447, 449, 451, 453, 455, 457, 459, 461, 463, 465, 467, 469, 471, 473, 475, 477, 479, 481, 483, 485 and 487;
 - ii) has an amino acid sequence encoded by a nucleic acid sequence recited in any one of SEQ ID Nos: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 92a, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200,

202, 204, 206, 208, 210, 212, 214 and 216, or has an amino acid sequence encoded by a gene identified from an EST recited in any one of these SEQ ID Nos;

iii) is a fragment of a polypeptide according to i) or ii), provided that said fragment retains a biological activity possessed by the full length polypeptide of i) or ii), or has an antigenic determinant in common with the polypeptide of i) or ii); or

iv) is a functional equivalent of a polypeptide of i), ii) or (iii).

for use in the diagnosis or therapy of the disease or abnormal physiological condition that is affected by hypoxia, such as cancer, ischaemic conditions, reperfusion injury, retinopathy, neonatal stress, preeclampsia, atherosclerosis, inflammatory conditions, wound healing, tumourigenesis, angiogenesis, apoptosis, inflammation, erythropoiesis, hair loss, or the biological response to hypoxia conditions, including processes such as glycolysis, gluconeogenesis, glucose transportation, catecholamine synthesis, iron transport and nitric oxide synthesis.

36. A purified and isolated nucleic acid molecule that encodes a polypeptide as recited in claim 35, for use in the diagnosis or therapy of for use in the diagnosis or therapy of a disease or abnormal physiological condition that is affected by hypoxia, such as cancer, ischaemic conditions, reperfusion injury, retinopathy, neonatal stress, preeclampsia, atherosclerosis, inflammatory conditions, wound healing, tumourigenesis, angiogenesis, apoptosis, inflammation, erythropoiesis or hair loss.

37. A purified nucleic acid molecule as recited in claim 36, which consists of the nucleic acid sequence as recited in any one of SEQ ID Nos.: 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, 268, 270, 272, 274, 276, 278, 280, 282, 284, 286, 288, 290, 292, 294, 296, 298, 300, 302, 304, 306, 308, 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330, 332, 334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 370, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 392, 394, 396, 398, 400, 402, 404, 406, 408, 410, 412, 414, 416, 418, 420, 422, 424, 426, 428, 430, 432, 434, 436, 438, 440, 442, 444, 446, 448, 450, 452, 454, 456, 458, 460, 462, 464, 466, 468, 470, 472, 474, 476, 478, 480, 482, 484, 486 and 488, or which is a redundant equivalent or fragment thereof, for use in the diagnosis or therapy of a disease or abnormal physiological condition that is affected by hypoxia, such as cancer, ischaemic conditions, reperfusion injury, retinopathy, neonatal stress, preeclampsia, atherosclerosis, inflammatory conditions, wound healing, tumourigenesis, angiogenesis, apoptosis, inflammation, erythropoiesis or hair loss.

38. A purified nucleic acid molecule which hybridizes under high stringency conditions with a nucleic acid molecule as recited in claim 36 or claim 37, for use in the diagnosis or therapy of a disease or

abnormal physiological condition that is affected by hypoxia, such as cancer, ischaemic conditions, reperfusion injury, retinopathy, neonatal stress, preeclampsia, atherosclerosis, inflammatory conditions, wound healing, tumourigenesis, angiogenesis, apoptosis, inflammation, erythropoiesis, or hair loss.

- 5 39. A vector comprising a nucleic acid molecule as recited in any one of claims 36-38, for use in the diagnosis or therapy of a disease or abnormal physiological condition that is affected by hypoxia, such as cancer, ischaemic conditions, reperfusion injury, retinopathy, neonatal stress, preeclampsia, atherosclerosis, inflammatory conditions, wound healing, tumourigenesis, angiogenesis, apoptosis, inflammation, erythropoiesis, or hair loss.
- 10 40. A ligand which binds specifically to, and which preferably inhibits the hypoxia-induced activity of, a polypeptide as recited in claim 35, for use in the diagnosis or therapy of tumourigenesis, angiogenesis, apoptosis, the biological response to hypoxia conditions, or a hypoxic-associated pathology.
41. A pharmaceutical composition suitable for modulating hypoxia and/or ischaemia, comprising a therapeutically-effective amount of a polypeptide as recited in any one of claims 13-21 or 35, a nucleic acid molecule according to any one of claims 22-24 or 36-38, a vector according to claim 25 or 39, or
15 a ligand according to claim 30, 31 or 40, in conjunction with a pharmaceutically-acceptable carrier.
42. A pharmaceutical composition according to claim 41, wherein said pharmaceutically-acceptable carrier is a liposome.
43. A vaccine composition comprising a polypeptide as recited in any one of claims 13-21 or 35, a nucleic acid molecule as recited in any one of claims 22-24 or 36-38, or a vector according to claim 25 or 39.
20
44. A method of treating a disease in a patient in need of such treatment by administering to a patient a therapeutically effective amount of a polypeptide as recited in any one of claims 13-21 or 35, an antagonist of said polypeptide, or a nucleic acid molecule as recited in any one of claims 22-24 or 36-38.
- 25 45. A method of regulating tumourigenesis, angiogenesis, apoptosis, the biological response to hypoxia conditions, or a hypoxic-associated pathology in a patient in need of such treatment by administering to a patient a therapeutically effective amount of a polypeptide according to any one of claims 13-21 or 35, a nucleic acid molecule according to any one of claims 22-24 or 36-38, or a vector according to claim 25 or 39, or a ligand according to claim 30, 31 or 40 or a pharmaceutical composition according
30 to claim 41 or 42.

46. A method according to claim 45, wherein, for diseases in which the expression of the natural gene or the activity of the polypeptide is lower in a diseased patient when compared to the level of expression or activity in a healthy patient, the polypeptide, nucleic acid molecule, ligand, compound or composition administered to the patient is an agonist.
- 5 47. A method according to claim 45, wherein, for diseases in which the expression of the natural gene or activity of the polypeptide is higher in a diseased patient when compared to the level of expression or activity in a healthy patient, the polypeptide, nucleic acid molecule, vector, ligand, compound or composition administered to the patient is an antagonist.
48. A polypeptide according to any one of claims 13-21 or 35, a nucleic acid molecule according to any
10 one of claims 22-24 or 36-38, a vector according to claim 25 or 39, a ligand according to claim 30, 31 or 40 or a pharmaceutical composition according to claim 41 or 42, for use in the manufacture of a medicament for the treatment of a hypoxia-regulated condition.
49. A method of monitoring the therapeutic treatment of a disease or physiological condition in a patient, comprising monitoring over a period of time the level of expression or activity of polypeptide
15 according to any one of claims 13-21 or 35, a nucleic acid molecule according to any one of claims 22-24 or 36-38, in tissue from said patient, wherein altering said level of expression or activity over the period of time towards a control level is indicative of regression of said disease.
50. A method of providing a hypoxia regulating gene, an apoptotic or an angiogenesis regulating gene by administering directly to a patient in need of such therapy an expressible vector comprising expression
20 control sequences operably linked to one or more of the nucleic acid molecules recited in claims 22-24 or 36-38.
51. A method of diagnosing a hypoxia-regulated condition in a patient, comprising assessing the level of expression of a natural gene encoding a polypeptide according to any one of claims 13-21 or 35, or assessing the activity of such a polypeptide, in tissue from said patient and comparing said level of
25 expression or activity to a control level, wherein a level that is different to said control level is indicative of the hypoxia-related condition.
52. A method according to claim 51 that is carried out *in vitro*.
53. A method according to claim 51 or claim 52, which comprises the steps of: (a) contacting a ligand according to claim 30, 31 or 40 with a biological sample under conditions suitable for the formation of
30 a ligand-polypeptide complex; and (b) detecting said complex.
54. A method according to claim 51 or claim 52, comprising the steps of:

- a) contacting a sample of tissue from the patient with a nucleic acid probe under stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule according to any one of claims 22-24 or 36-38 and the probe;
- b) contacting a control sample with said probe under the same conditions used in step a); and
- 5 c) detecting the presence of hybrid complexes in said samples;
- wherein detection of levels of the hybrid complex in the patient sample that differ from levels of the hybrid complex in the control sample is indicative of the hypoxia-related condition.

55. A method according to claim 51 or claim 52, comprising the steps of:

- a) contacting a sample of nucleic acid from tissue of the patient with a nucleic acid primer under
10 stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule according to any one of claims 22-24 or 36-38 and the primer;
- b) contacting a control sample with said primer under the same conditions used in step a);
- c) amplifying the sampled nucleic acid; and
- d) detecting the level of amplified nucleic acid from both patient and control samples;
- 15 wherein detection of levels of the amplified nucleic acid in the patient sample that differ significantly from levels of the amplified nucleic acid in the control sample is indicative of the hypoxia-related condition.

56. A method according to claim 51 or claim 52, comprising the steps of:

- a) obtaining a tissue sample from a patient being tested for the hypoxia-related condition;
- 20 b) isolating a nucleic acid molecule according to any one of claims 22-24 or 36-38 from said tissue sample; and
- c) diagnosing the patient for disease by detecting the presence of a mutation which is associated with the hypoxia-related condition in the nucleic acid molecule as an indication of the hypoxia-related condition.

25 57. The method of claim 56, further comprising amplifying the nucleic acid molecule to form an amplified product and detecting the presence or absence of a mutation in the amplified product.

58. A method according to any one of claims 49-57, wherein said disease is cancer, ischaemic conditions, reperfusion injury, retinopathy, neonatal stress, preeclampsia, atherosclerosis, inflammatory

conditions, wound healing, tumourigenesis, angiogenesis, apoptosis, inflammation, erythropoiesis, or hair loss.

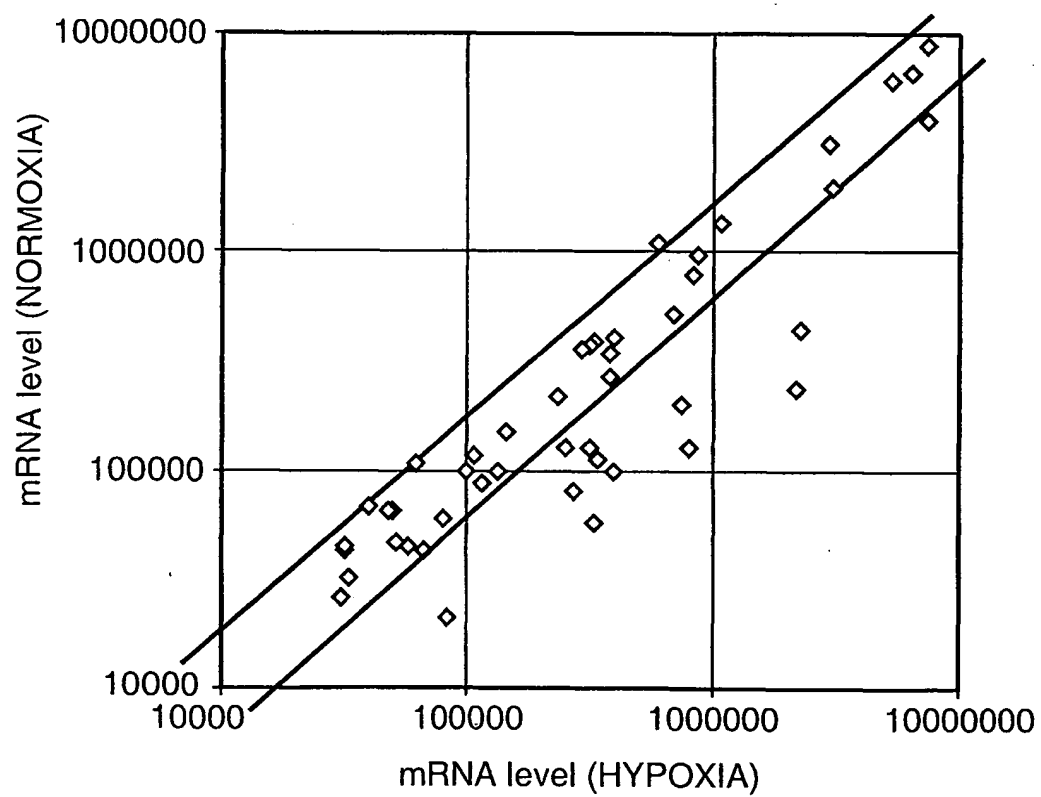
59. A method according to claim 58, wherein said hypoxia or ischaemia-related tissue damage is due to a disorder of the cerebral, coronary or peripheral circulation.
- 5 60. A method according to any one of claims 49, and 54-59, wherein the tissue is a cancer tissue.
61. A method for the identification of a compound that is effective in the treatment and/or diagnosis of disease, comprising contacting a polypeptide according to any one of claims 13-21 or 35, a nucleic acid molecule according to any one of claims 22-24 or 36-38 with one or more compounds suspected of possessing binding affinity for said polypeptide or nucleic acid molecule, and selecting a compound
- 10 that binds specifically to said nucleic acid molecule or polypeptide.
62. A method for the identification of a compound that is effective in the treatment and/or diagnosis of disease, comprising contacting a cell or cell membrane preparation comprising a polypeptide according to any one of claims 13-21 or 35 or a nucleic acid molecule according to any one of claims 22-24 or 36-38 with one or more candidate compounds and detecting the degree of compound binding,
- 15 or the stimulation or inhibition of a functional response in said cell or cell membrane.
63. A compound identified or identifiable by a method according to claim 61 or claim 62.
64. A compound according to claim 63, which is a natural or modified substrate, an enzyme, a receptor, a small organic molecule, such as a small natural or synthetic organic molecule of up to 2000Da, preferably 800Da or less, a peptidomimetic, an inorganic molecule, a peptide, a polypeptide, an
- 20 antibody, or a structural or functional mimetics of any of these compounds.
65. A kit useful for diagnosing disease comprising a first container containing a nucleic acid probe that hybridises under stringent conditions with a nucleic acid molecule according to any one of claims 22-24 or 36-38; a second container containing primers useful for amplifying said nucleic acid molecule; and instructions for using the probe and primers for facilitating the diagnosis of disease.
- 25 66. The kit of claim 65, further comprising a third container holding an agent for digesting unhybridised RNA.
67. An array of at least two nucleic acid molecules, wherein each of said nucleic acid molecules either corresponds to the sequence of, is complementary to the sequence of, or hybridises specifically to a nucleic acid molecule according to any one of claims 22-24 or 36-38.

68. An array according to claim 67, which contains nucleic acid molecules that either correspond to the sequence of, are complementary to the sequence of, or hybridise specifically to at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 92a, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294 or 295 of the nucleic acid molecules implicated in a hypoxia-regulated condition as recited in claims 22-24 or 36-38.
69. An array according to any claim 67 or claim 68, wherein said nucleic acid molecules consist of between twelve and two thousand nucleotides.
70. An array of antibodies, comprising at least two different antibody species, wherein each antibody species is immunospecific with a polypeptide implicated in a hypoxia-regulated condition as recited in any one of claims 13-21 or 35.
71. An array of polypeptides, comprising at least two polypeptide species as recited in any one of claims 13-21 or 35, wherein each polypeptide species is implicated in a hypoxia-regulated condition, or is a functional equivalent variant or fragment thereof.
72. A kit comprising an array of nucleic acid molecules according to any one of claims 67-69.
73. A kit comprising one or more antibodies that bind to a polypeptide as recited in any one of claims 13-21 or 35; and a reagent useful for the detection of a binding reaction between said antibody and said polypeptide.
74. A transgenic or knockout non-human animal that has been transformed to express higher, lower or absent levels of a polypeptide according to any one of claims 13-21 or 35.

75. A method for screening for a compound effective to treat disease, by contacting a non-human transgenic animal according to claim 74 with a candidate compound and determining the effect of the compound on the disease or physiological condition of the animal.
76. A substantially purified polypeptide comprising the consensus sequence:
5 KAMVACYPGNGTGYVRHVDNPNNGDGRCITCIYYLNKNWDAKLHGGILRIFPEGKSFIADVEPI
FDRLLFFWSDRRNPHEVQPSYATRYAMTVWYFDAEERAEAKKK, or a variant thereof.
77. A substantially purified polypeptide according to claim 76, for use in the diagnosis or treatment of a hypoxia-related disease or condition.

1/96

FIG. 1



2/96

FIG. 2a Clone ID:p1A23
Metallothionein 2A

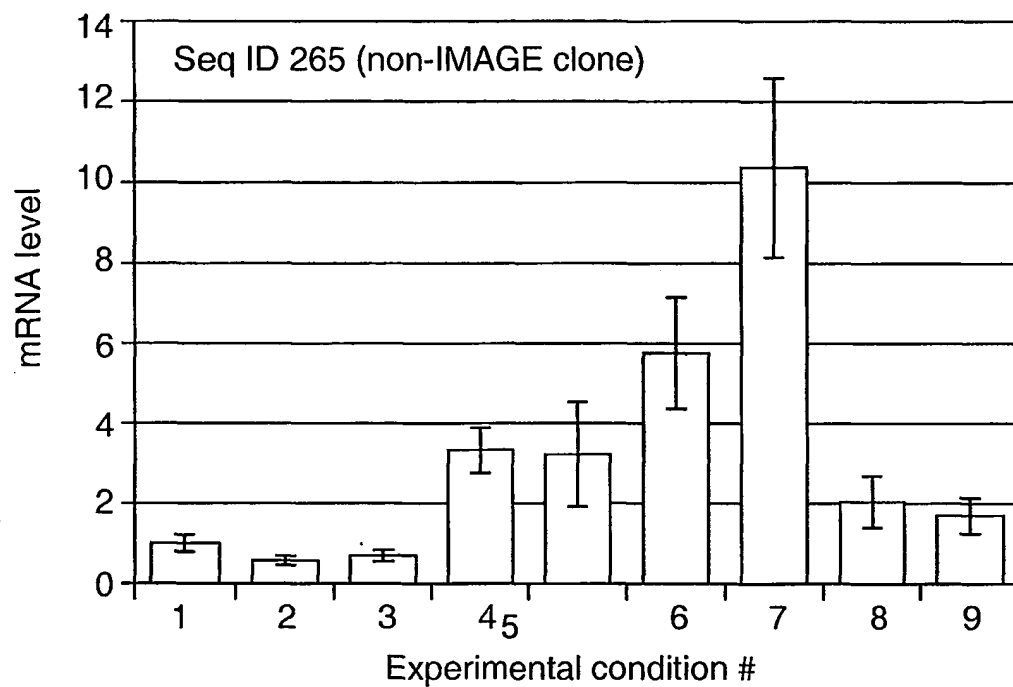
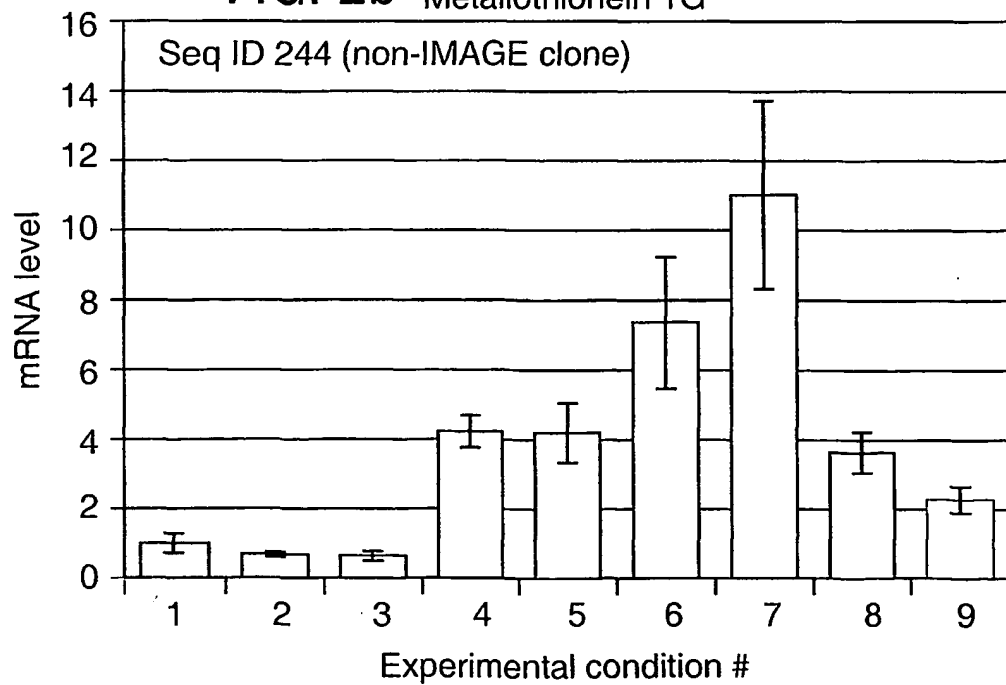


FIG. 2b Clone ID :p1B1
Metallothionein 1G



3/96

FIG. 2c

Clone ID: p1F6
Hypothetical protein hqp0376

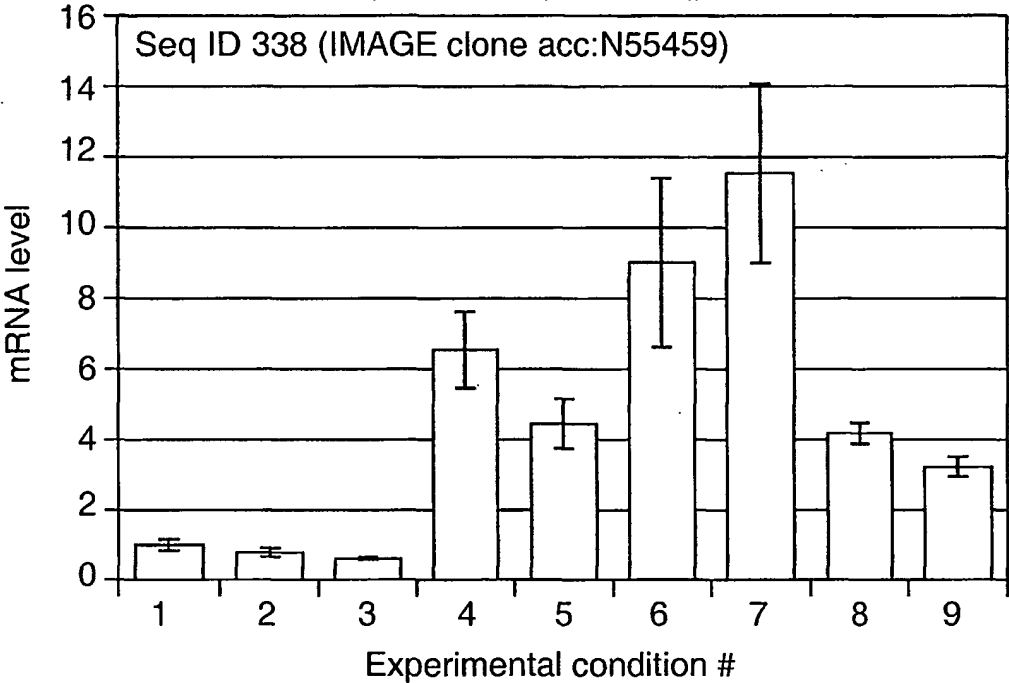
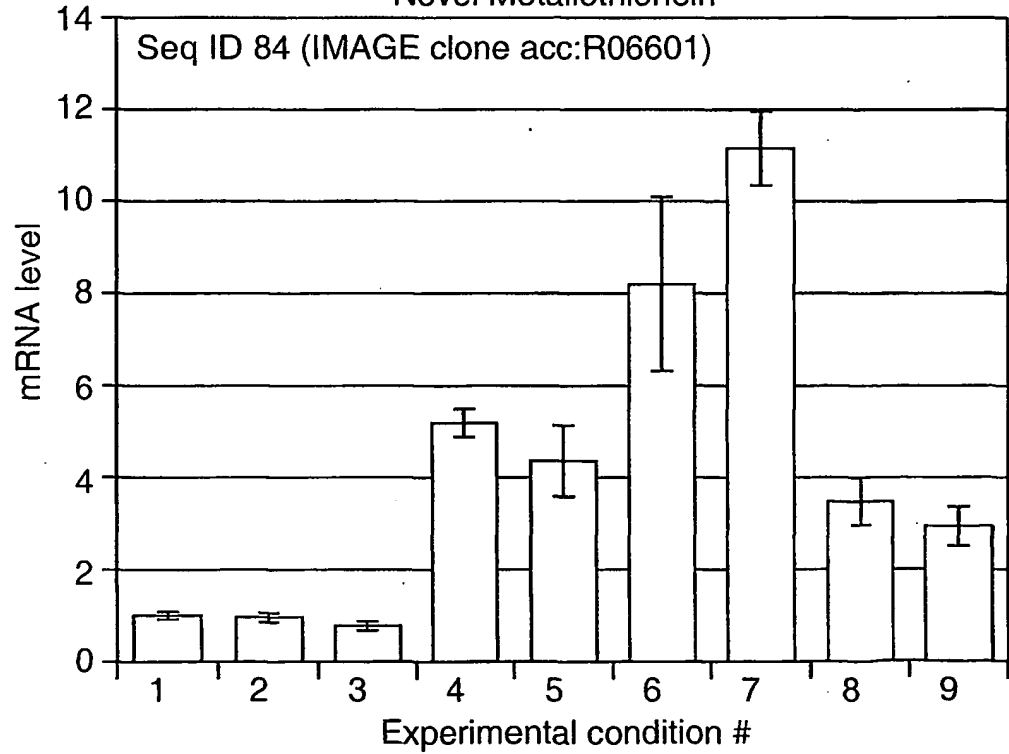


FIG. 2d

Clone ID: p1E7
Novel Metallothionein

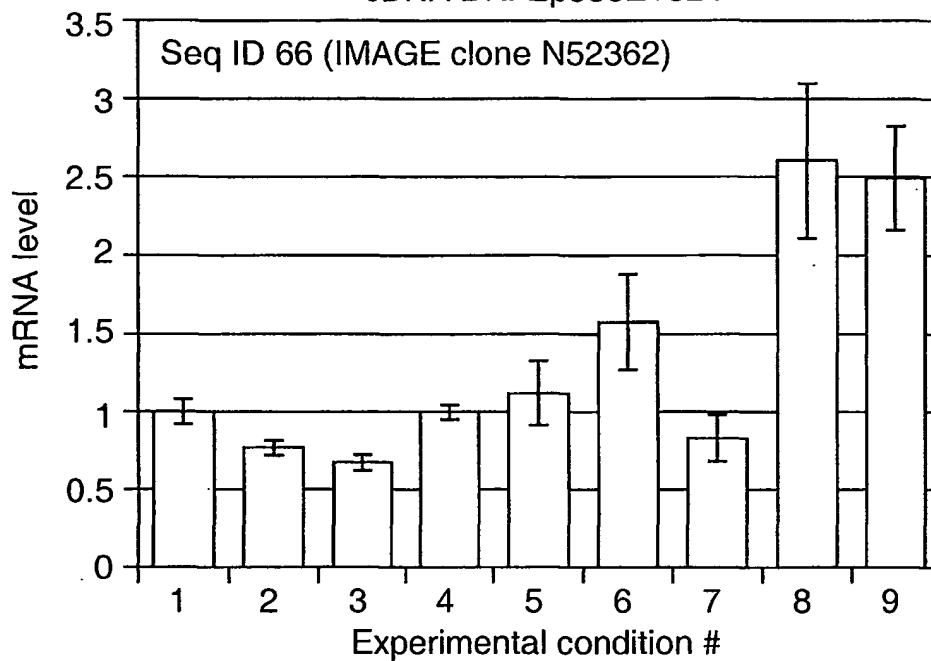


4/96

FIG.3a

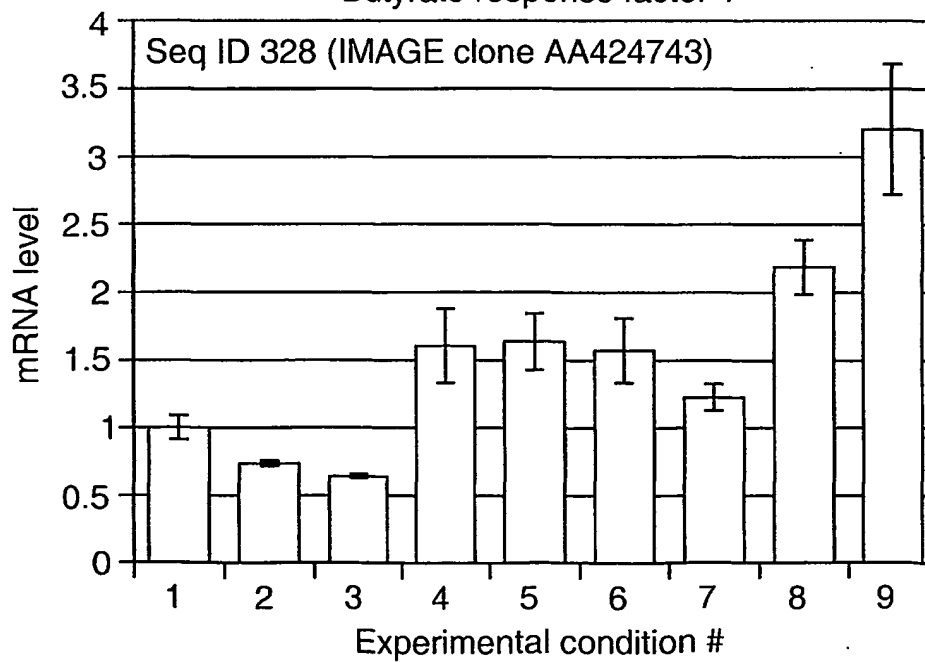
Clone p1E16

cDNA DKFZp586E1624

**FIG. 3b**

Clone p1F14

Butyrate response factor 1

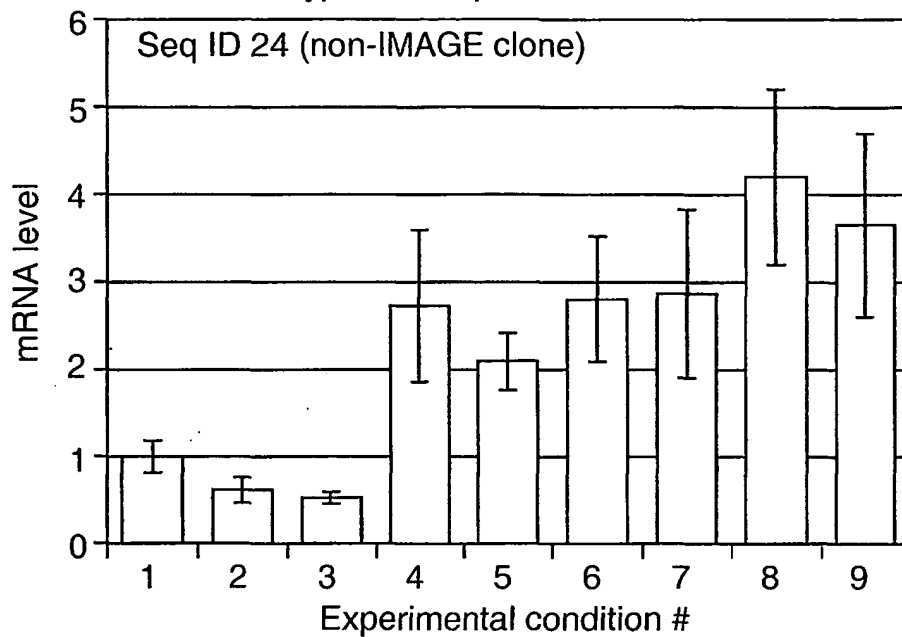


5/96

FIG .3c

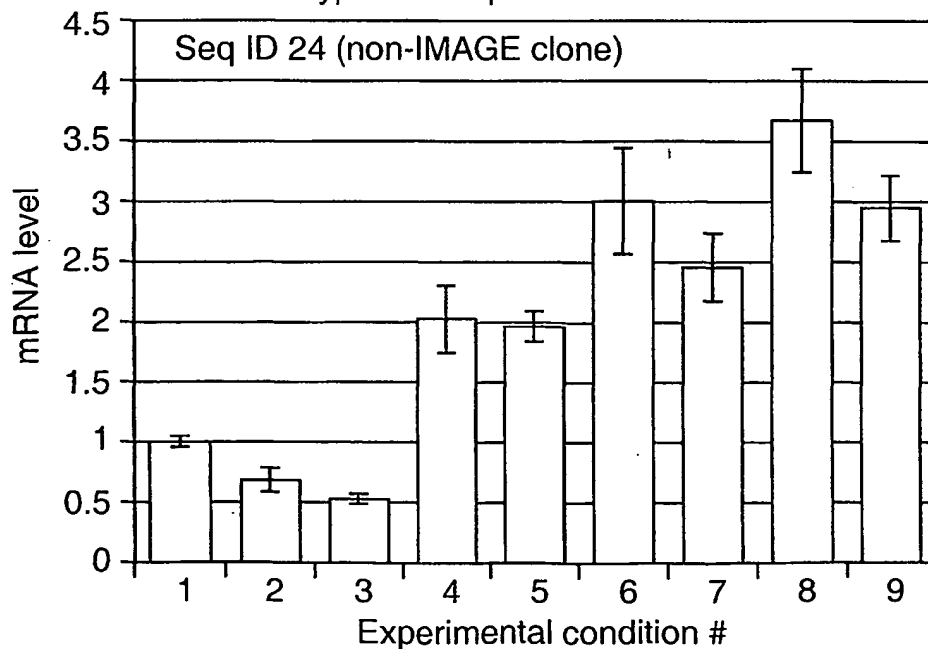
Clone p1D1

Hypothetical protein FLJ10134

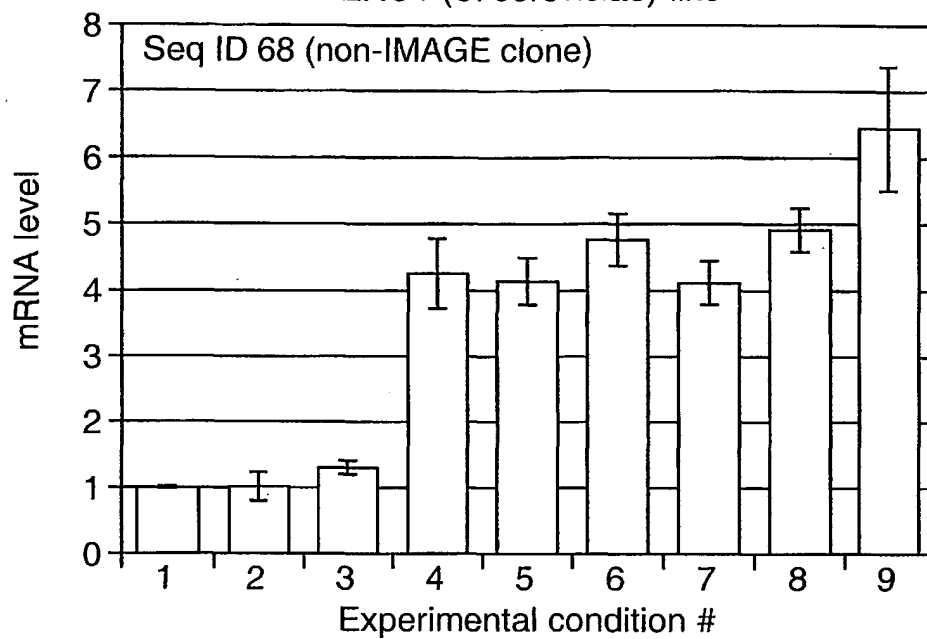
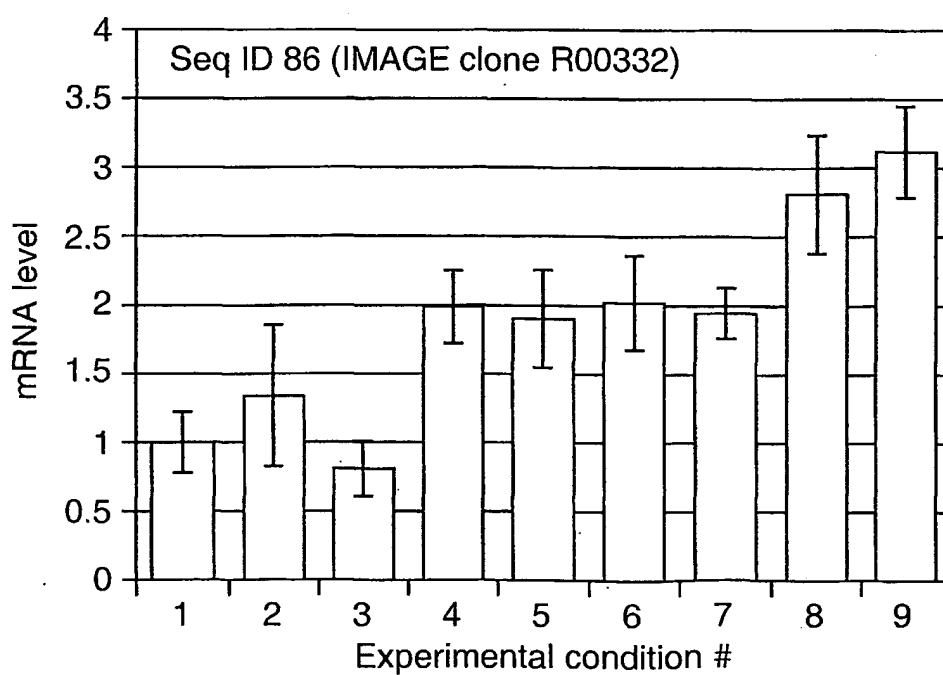
**FIG .3d**

Clone p1D2

Hypothetical protein FLJ10134



6/96

FIG .3eClone p1D6
ERO1 (*S. cerevisiae*)-like**FIG .3f**Clone p1E6
EGL nine (*C.elegans*) homolog

7/96

FIG .4

Clone p1115

Hypothetical protein CGI-117

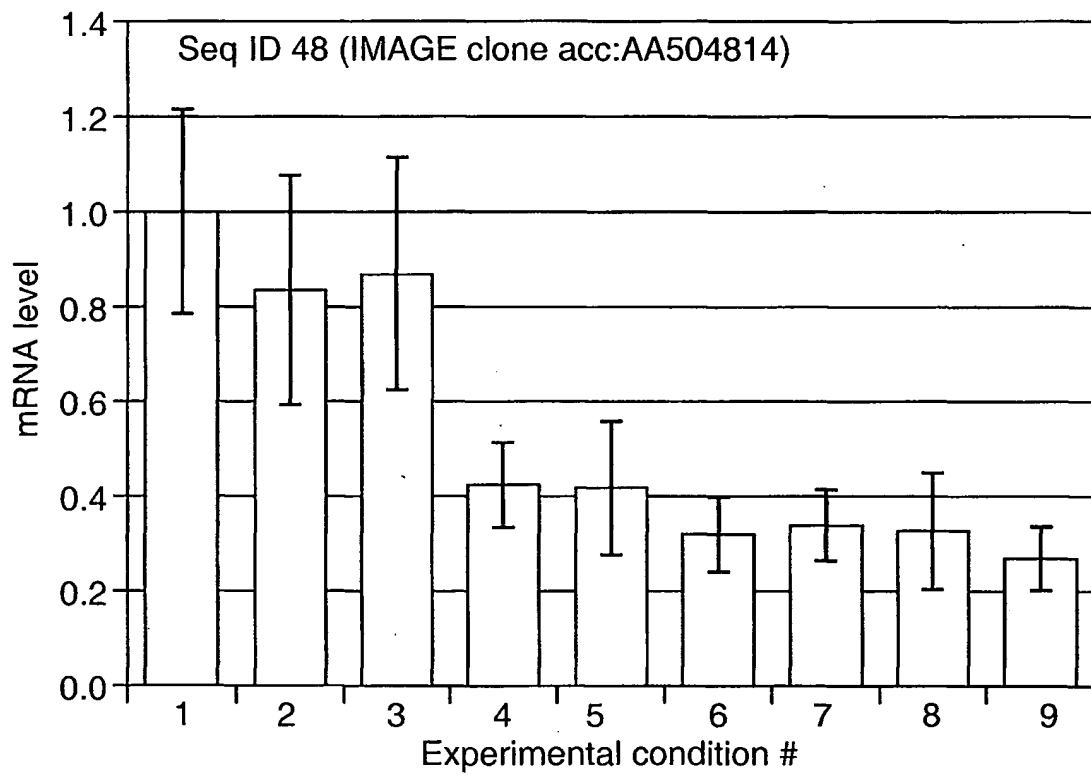
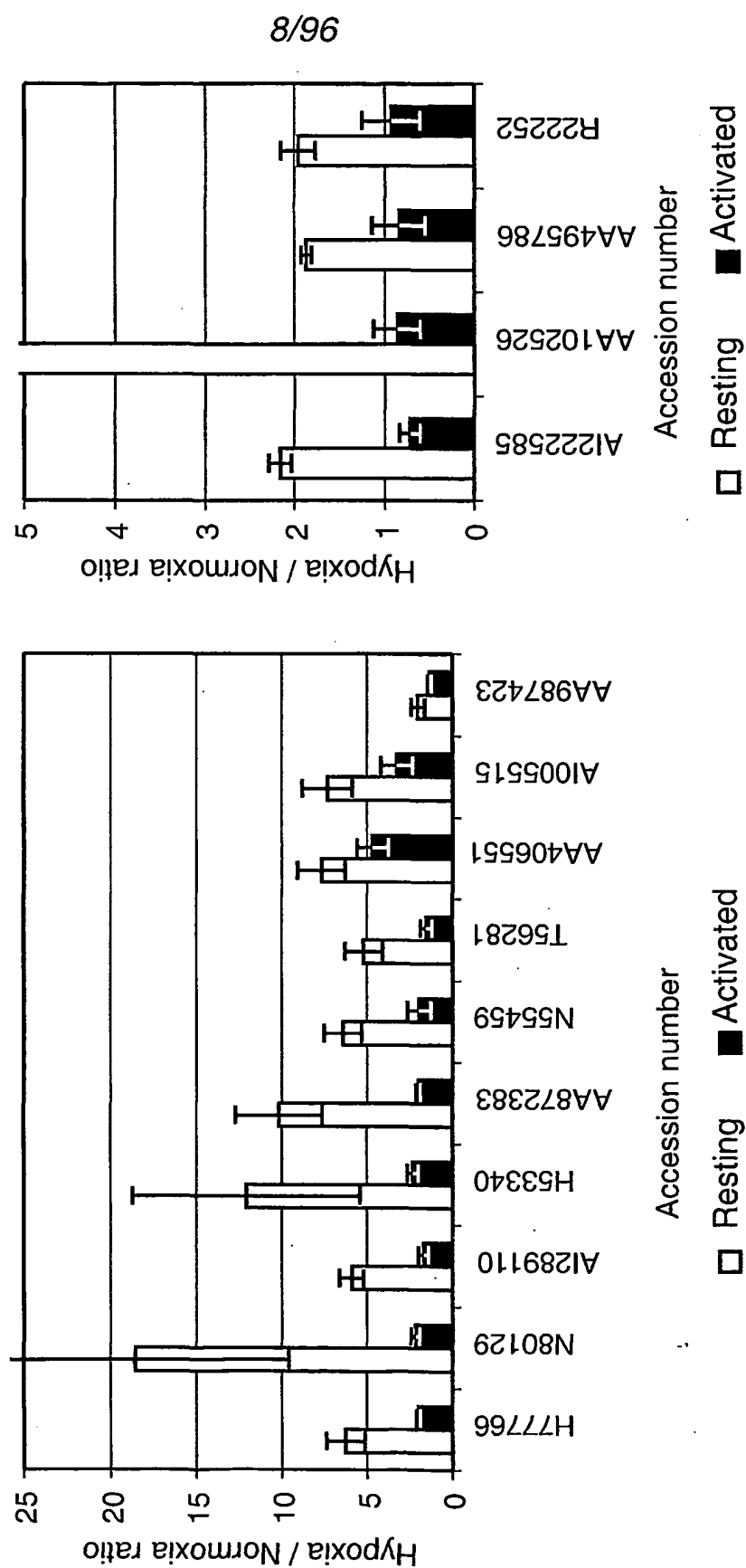
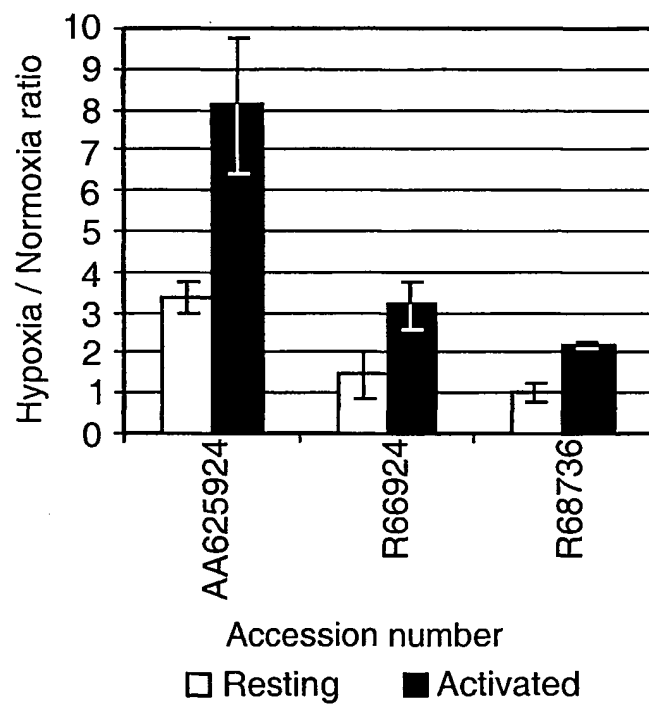


FIG. 5



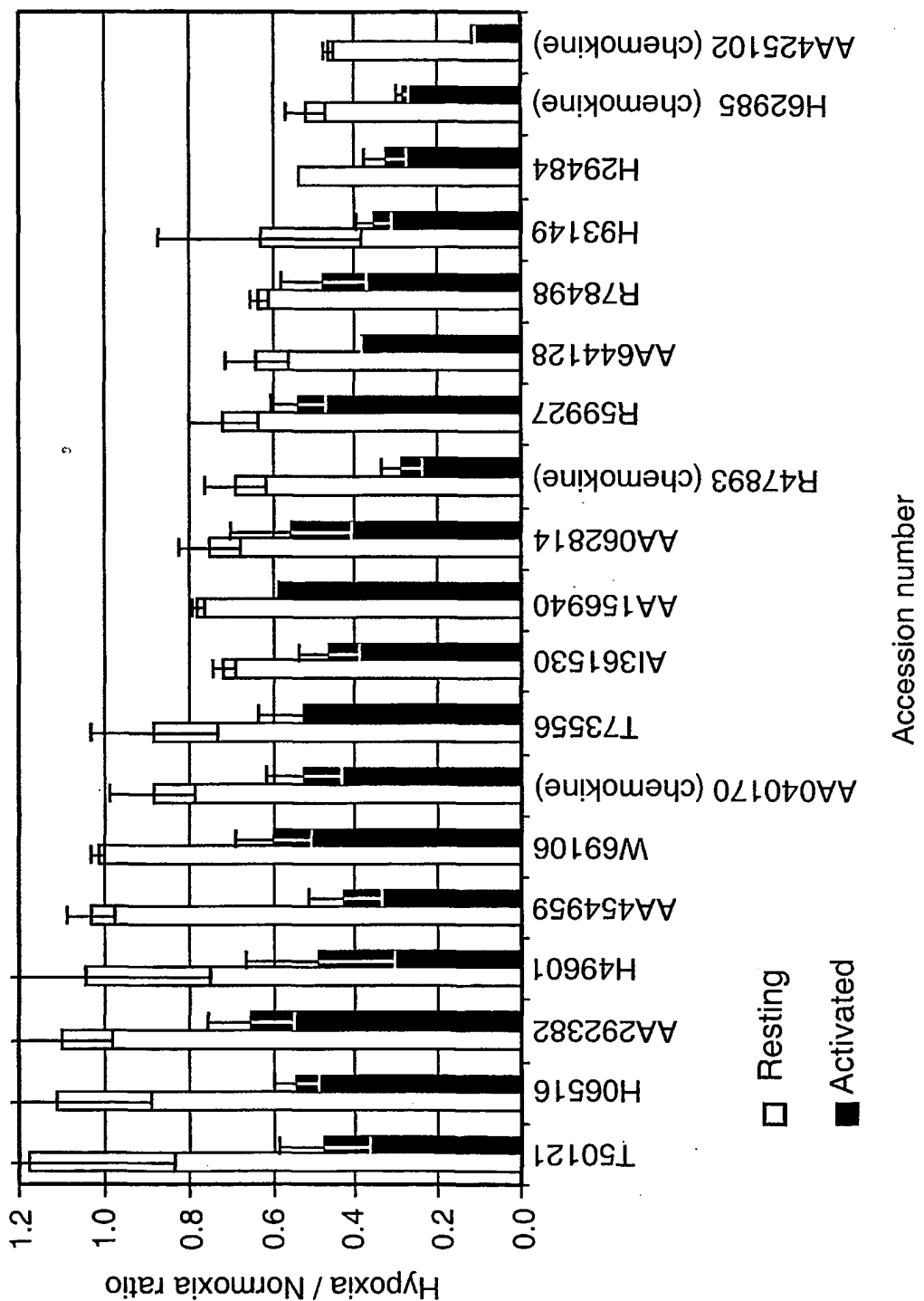
9/96

FIG. 6



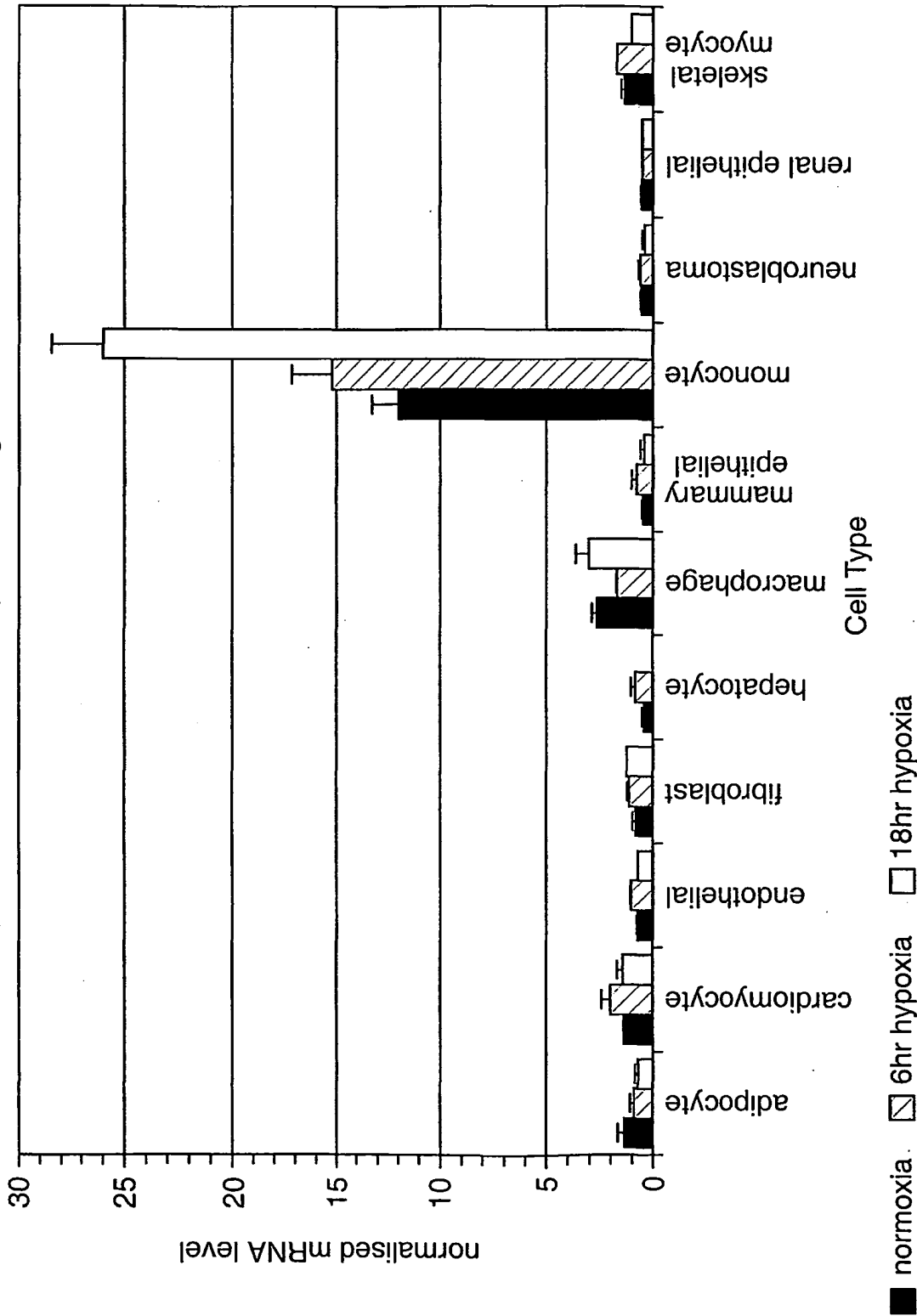
96/01

FIG. 7



11/96

FIG .8 p1123/ SeqID:476/ Ecotropic viral integration site 2A



12/96

FIG. 9a p1E9/ SeqID:80/ Novel PI-3-kinase adapter

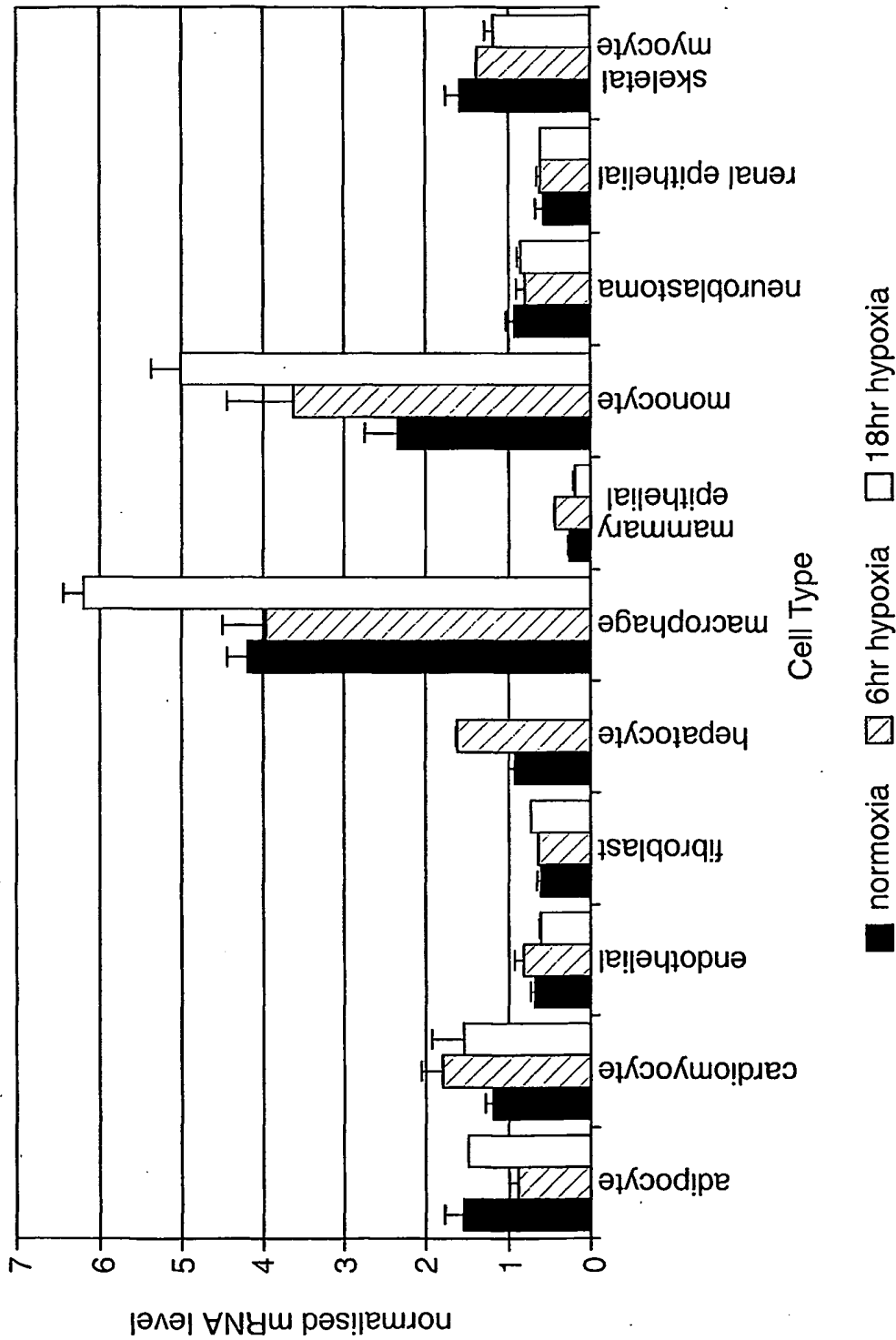
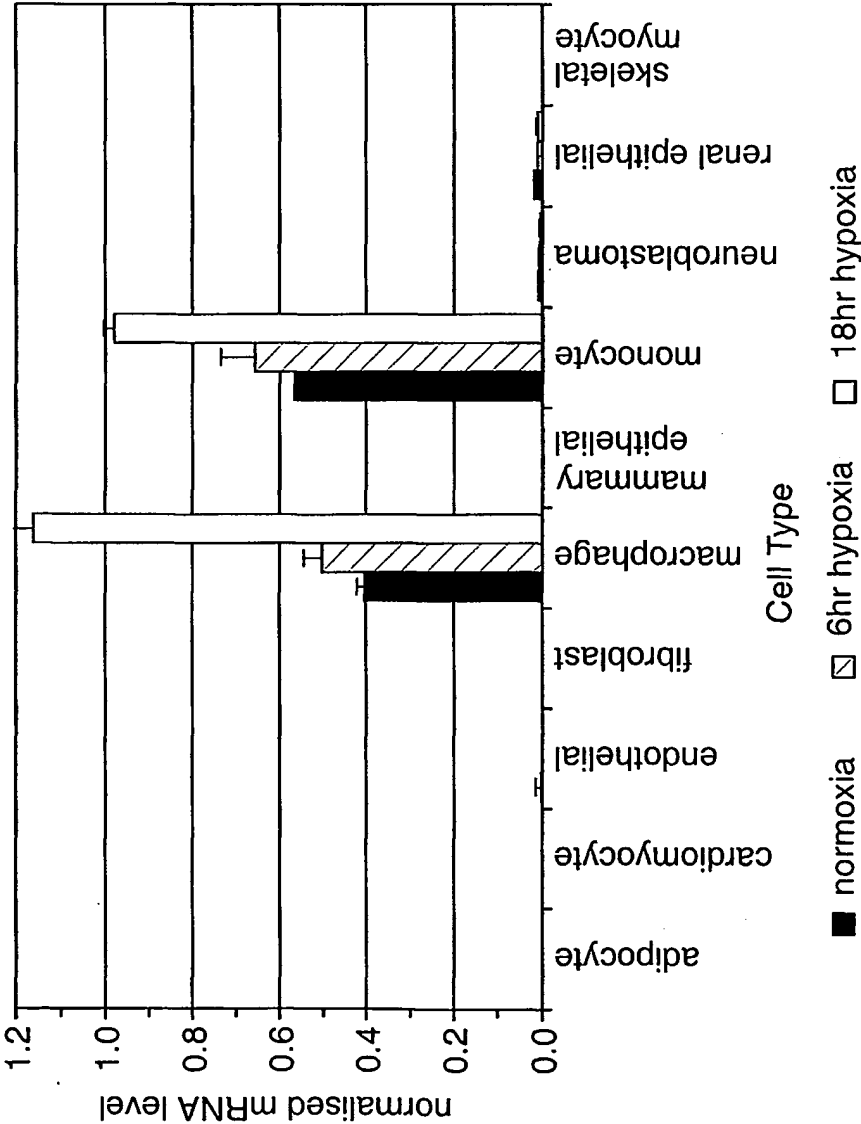


FIG. 9b p1E9/ SeqID:80/ Novel PI-3-kinase adapter



14/96

FIG. 9C p1N15/ IMAGE clone acc: R59598/ Syk

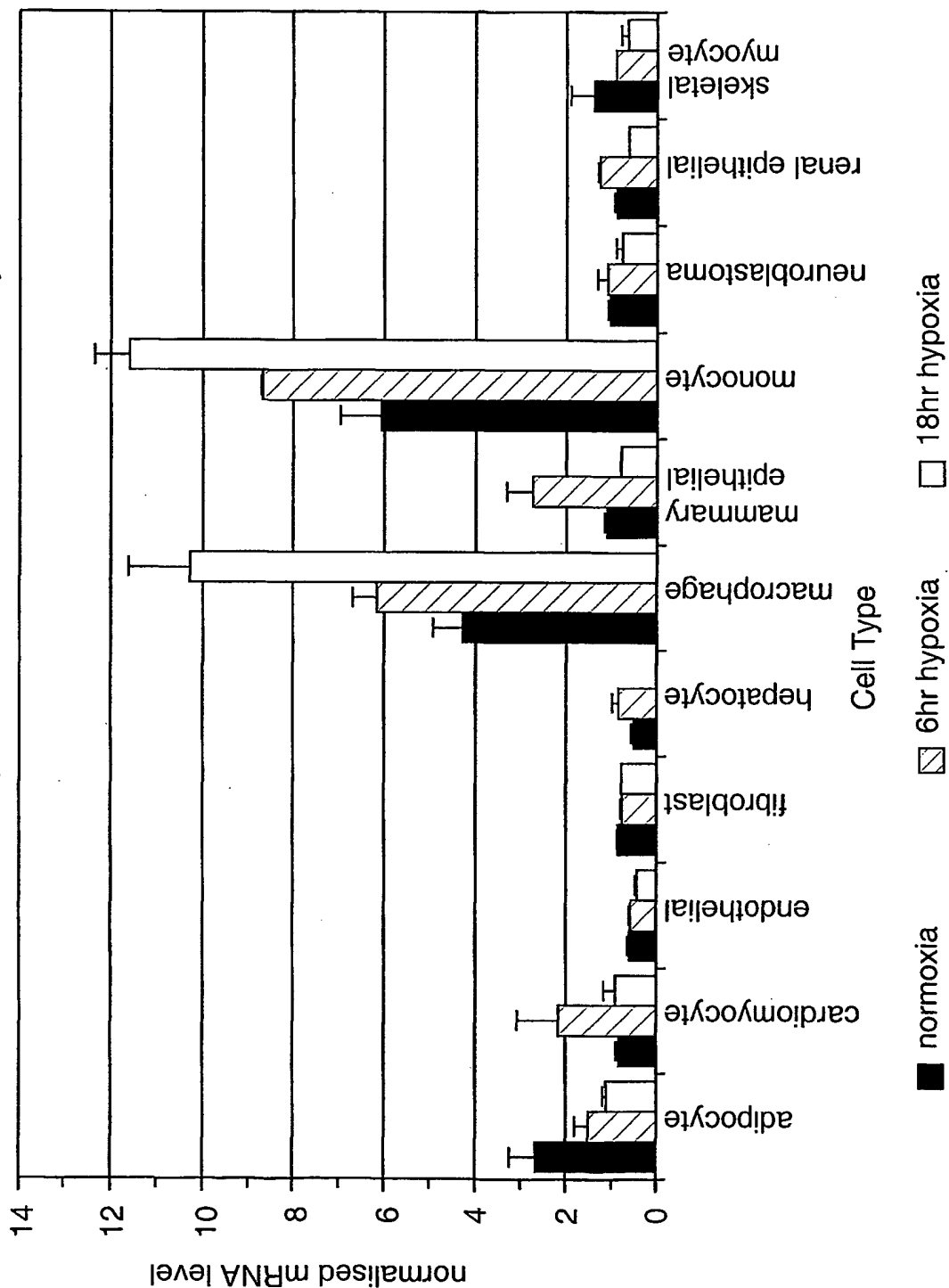


FIG .10 p1C10/ SeqID:376/ Regulator of G-protein signalling 1

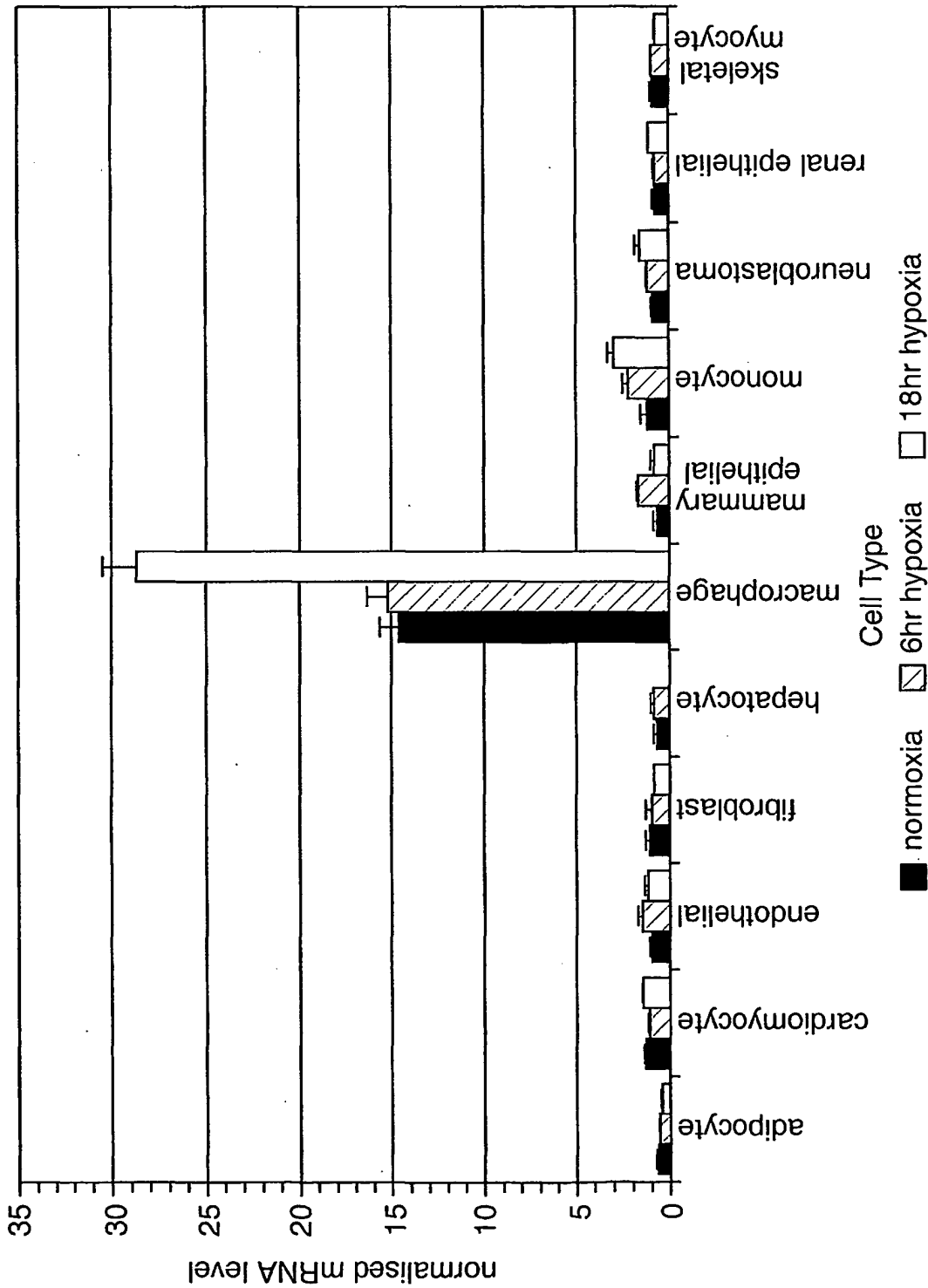
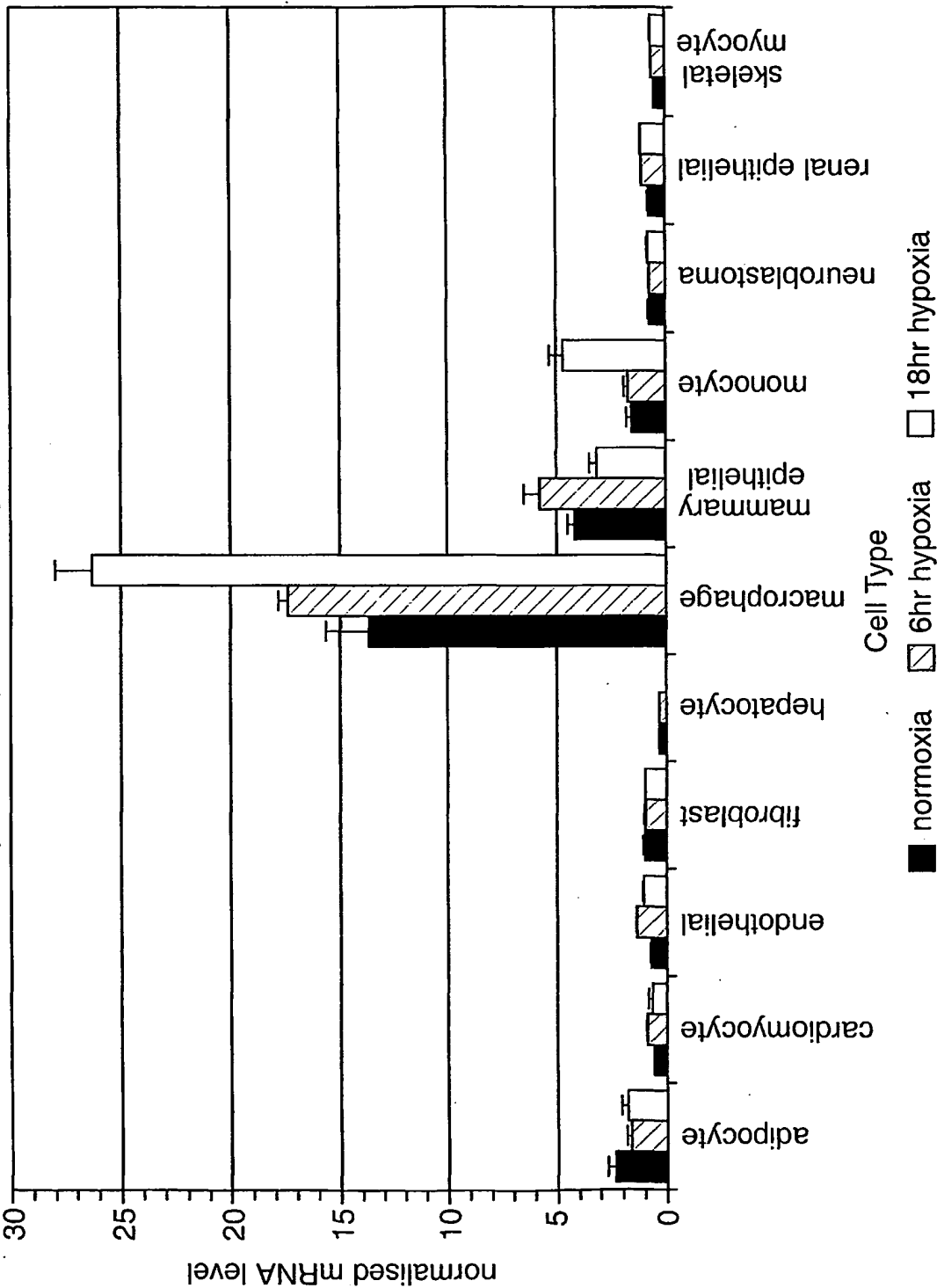
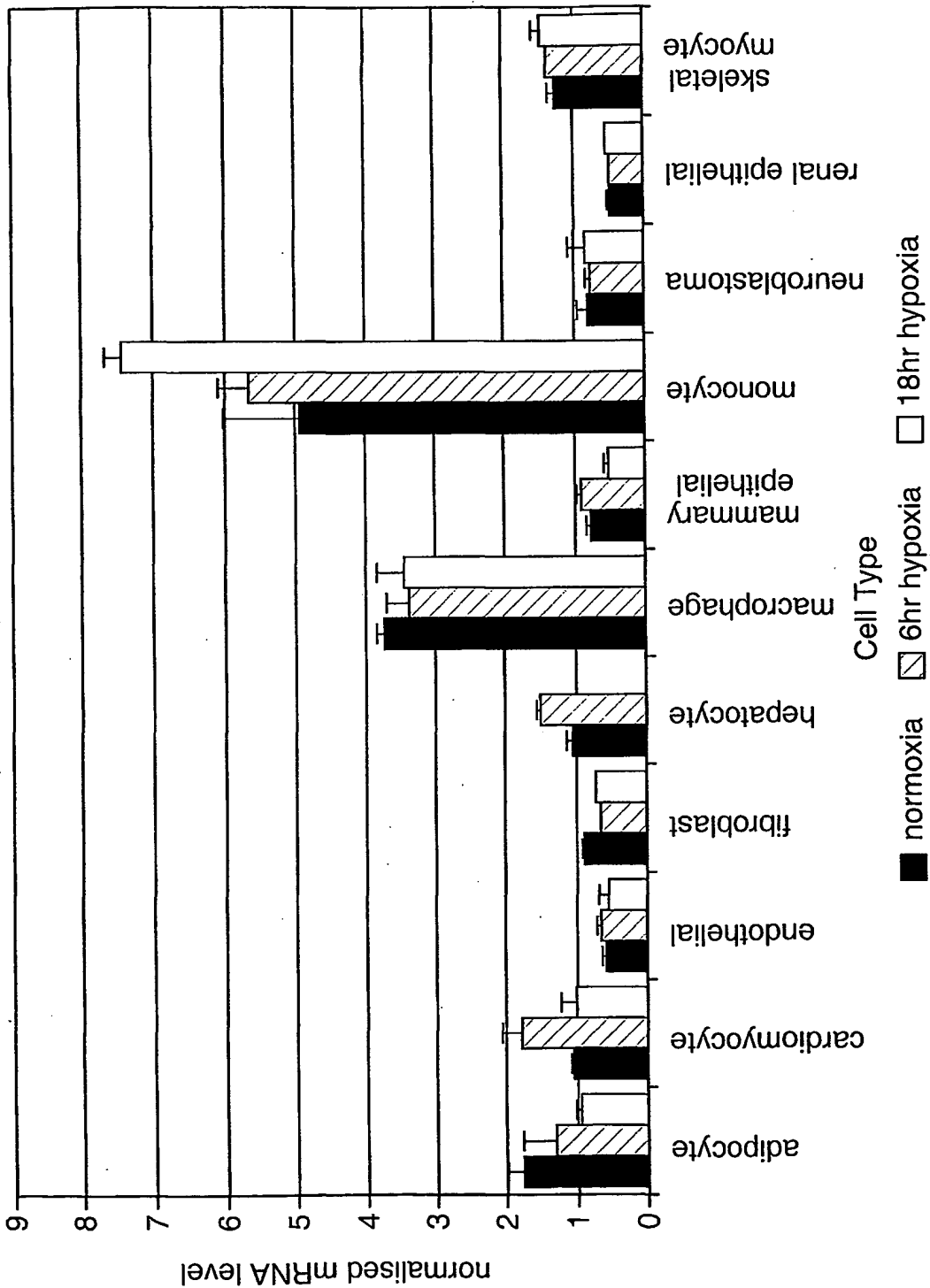


FIG 11 p1C19/ SeqID:390/ GM2 ganglioside activator protein



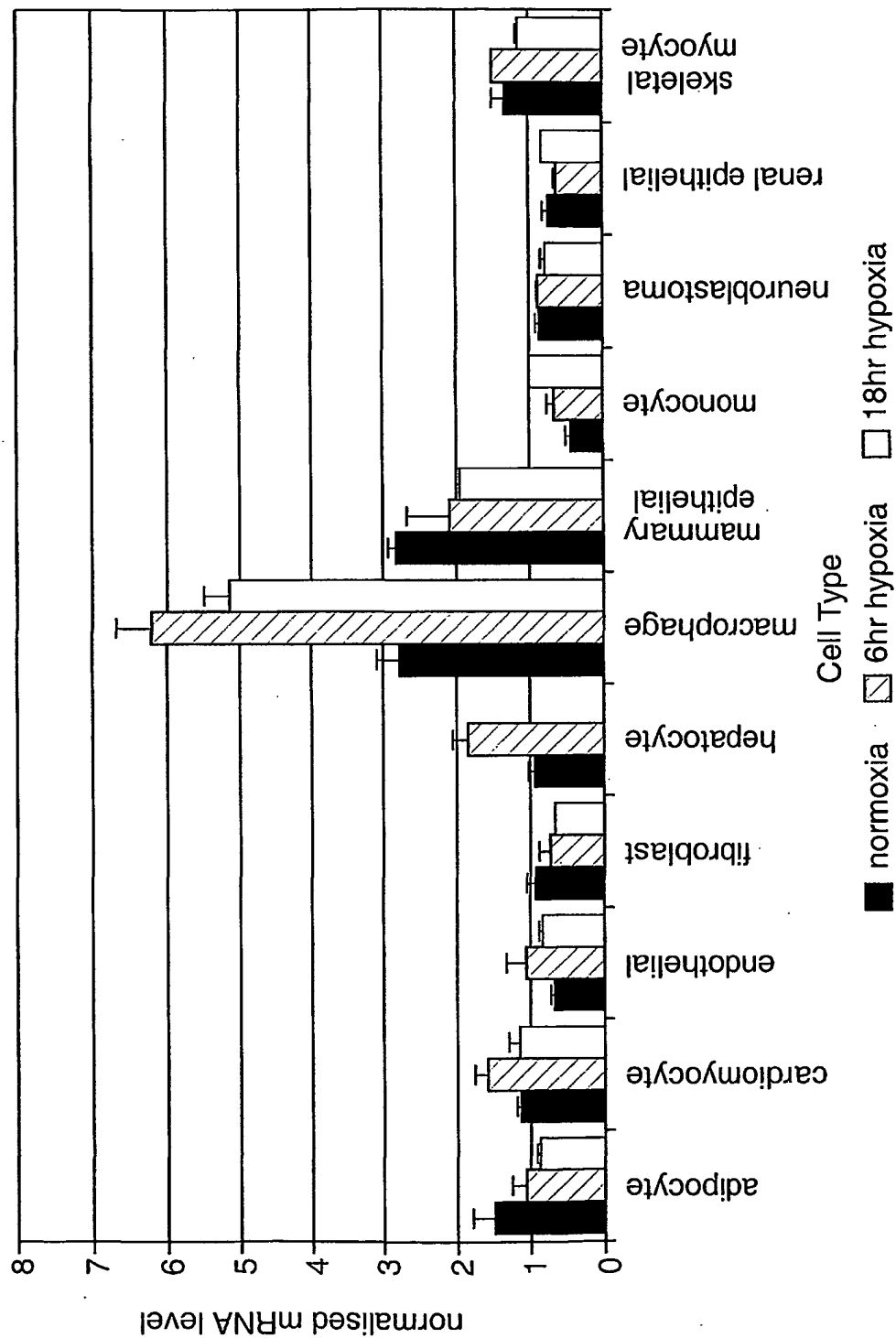
17/96

FIG .12 p1E13/ SeqID:22/ Hypothetical protein PRO0823



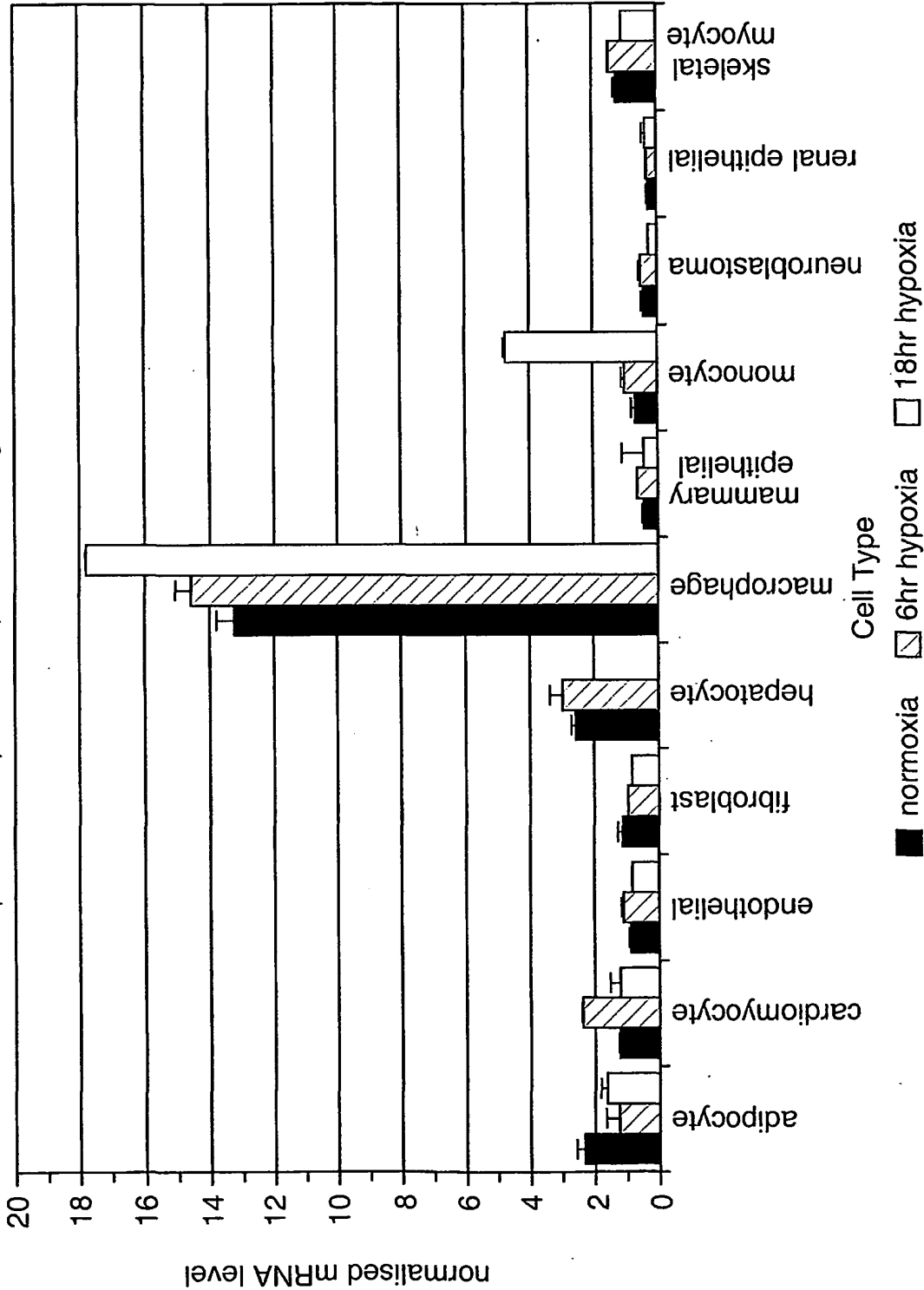
18/96

FIG. 13 p1F4/ SeqID:340/ CYP1



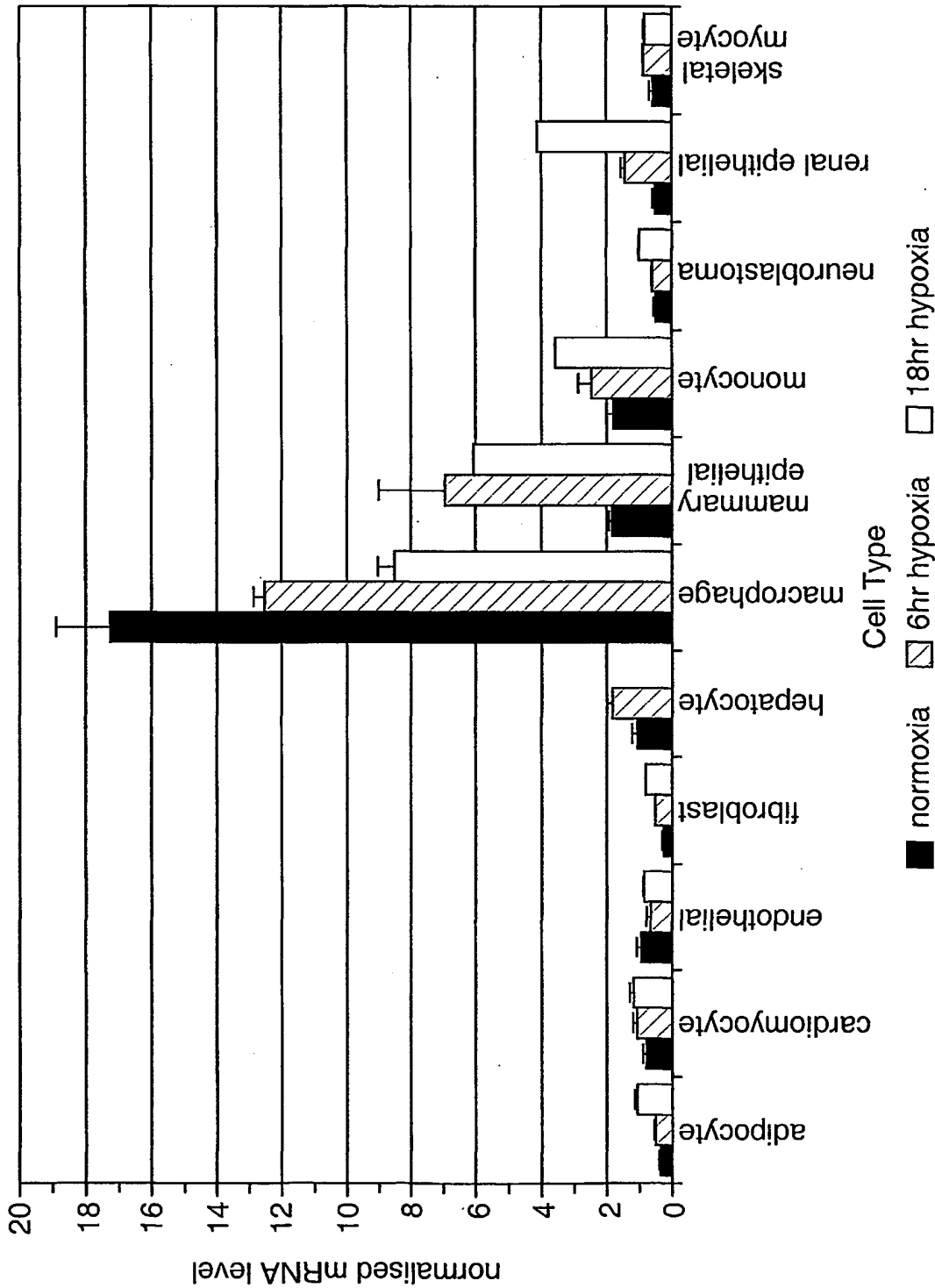
19/96

FIG. 14 p1K15/ SeqID:406/ Alpha-2-macroglobulin (seq ID: 405/406)



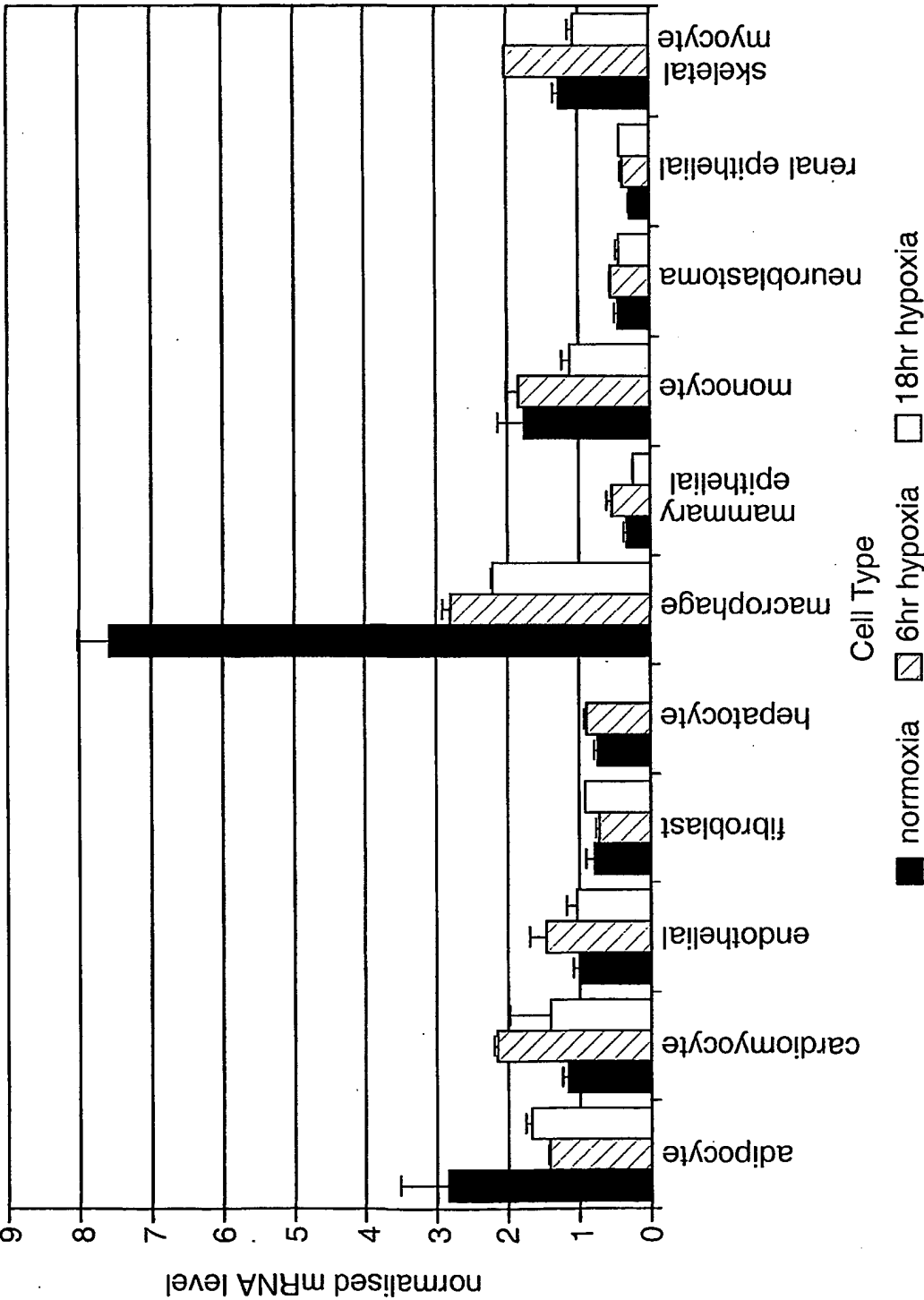
20/96

FIG. 15 p1B23/ SeqID:358/ interleukin 1 receptor antagonist



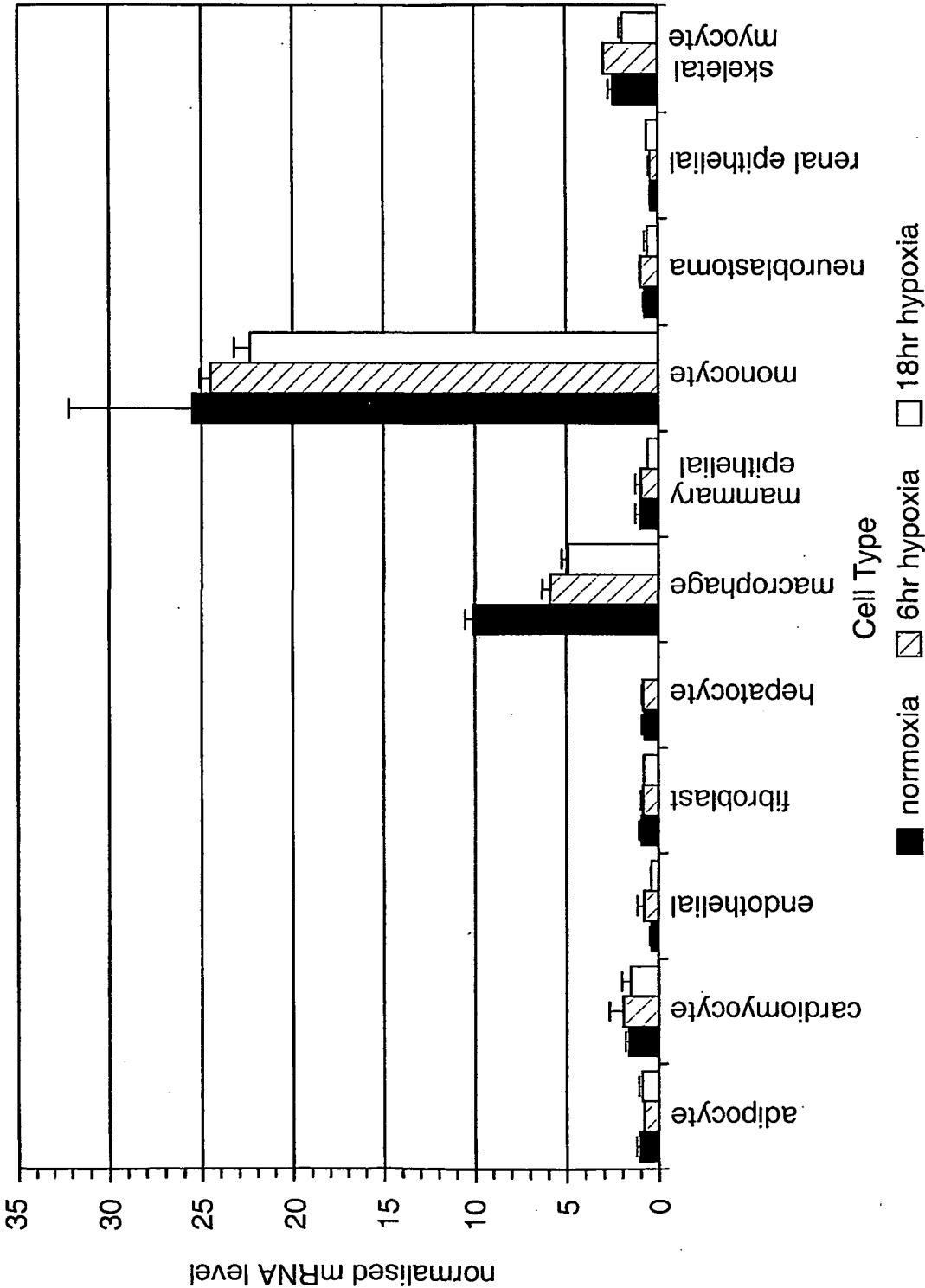
21/96

FIG. 16 p1120/ SeqID:470/ SCYA3L



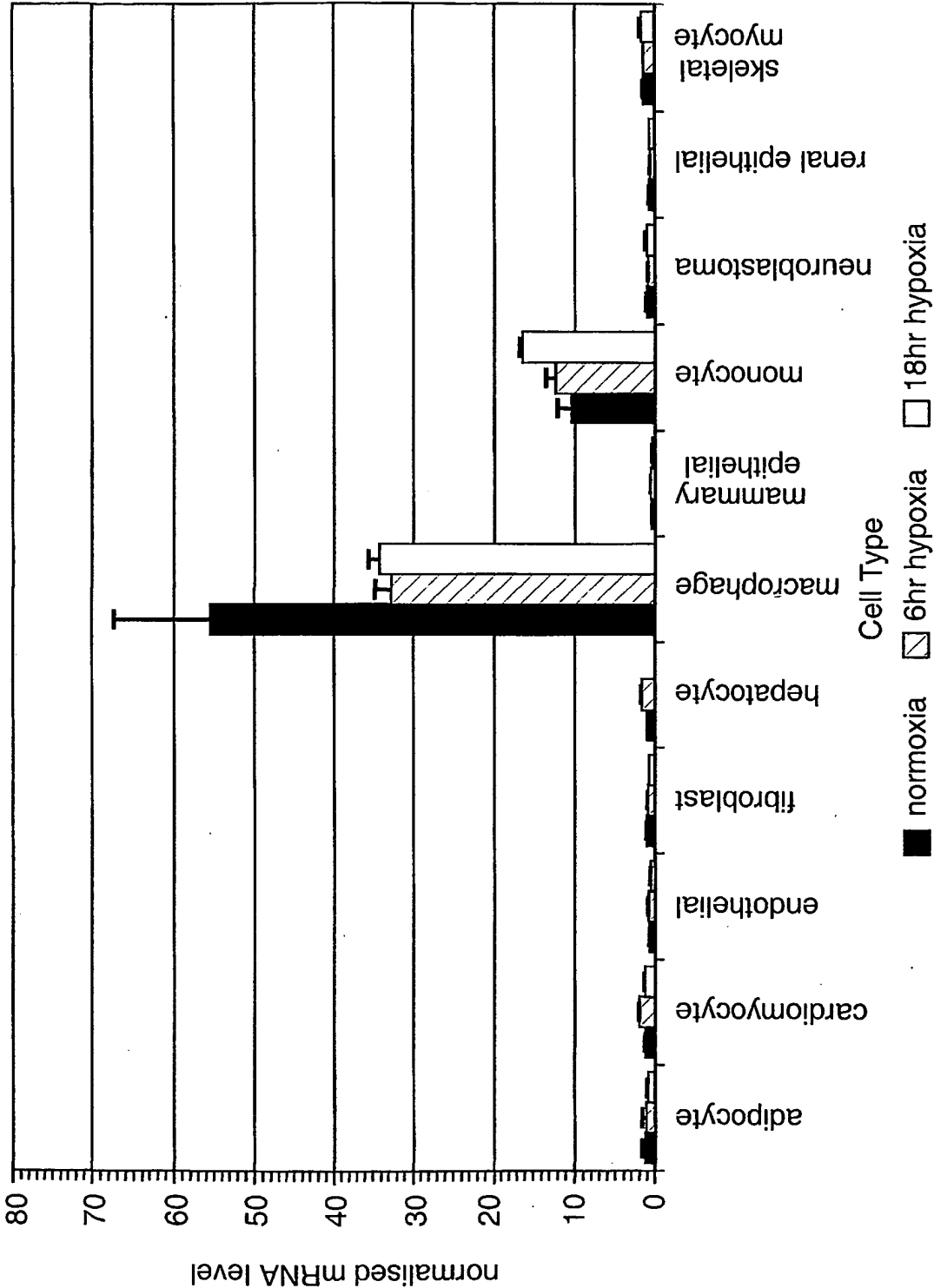
22/96

FIG. 17 p1K2/ SeqID:434/ CFFM4



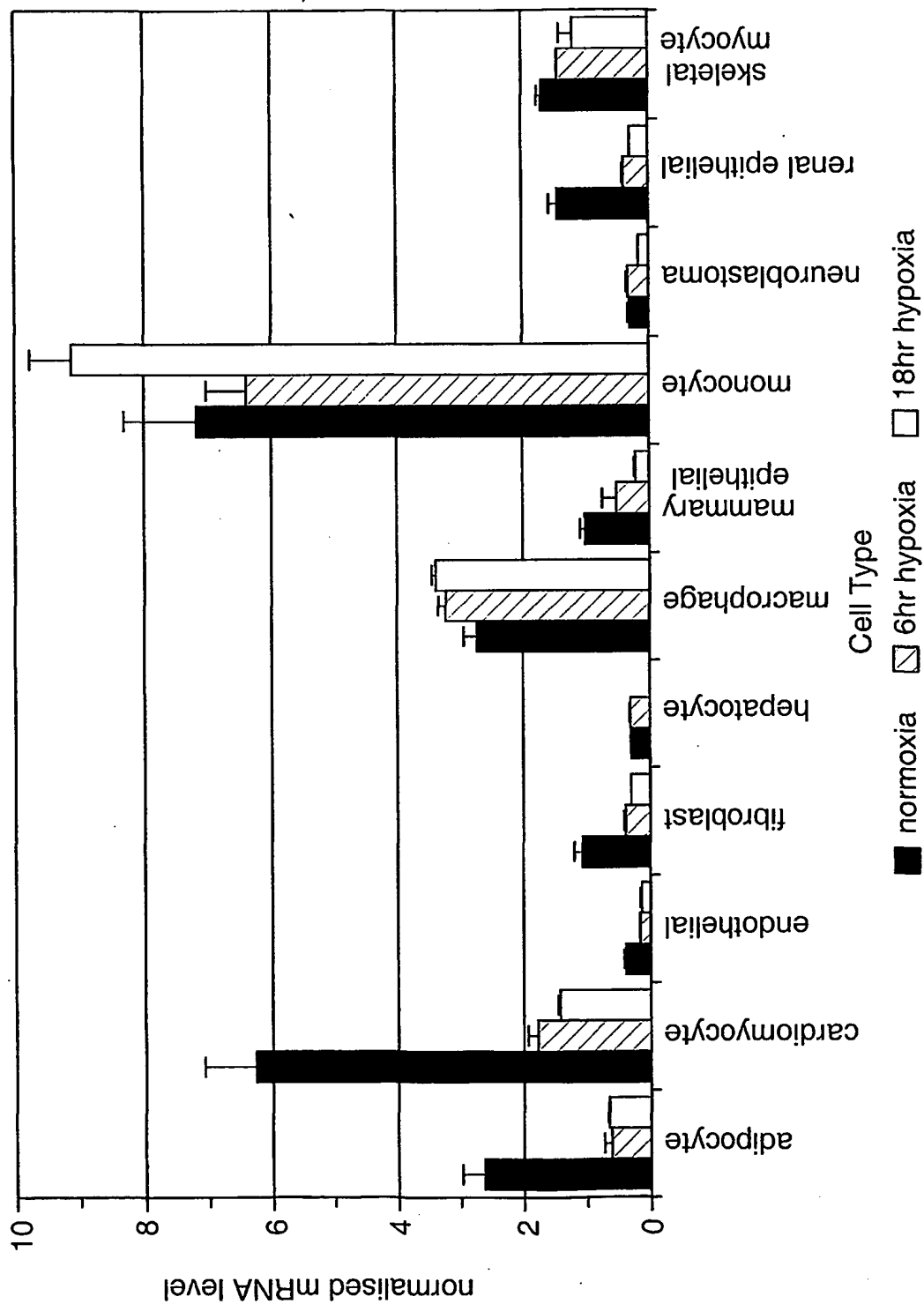
23/96

FIG. 18 p1K3/ SeqID:432/ Pleckstrin



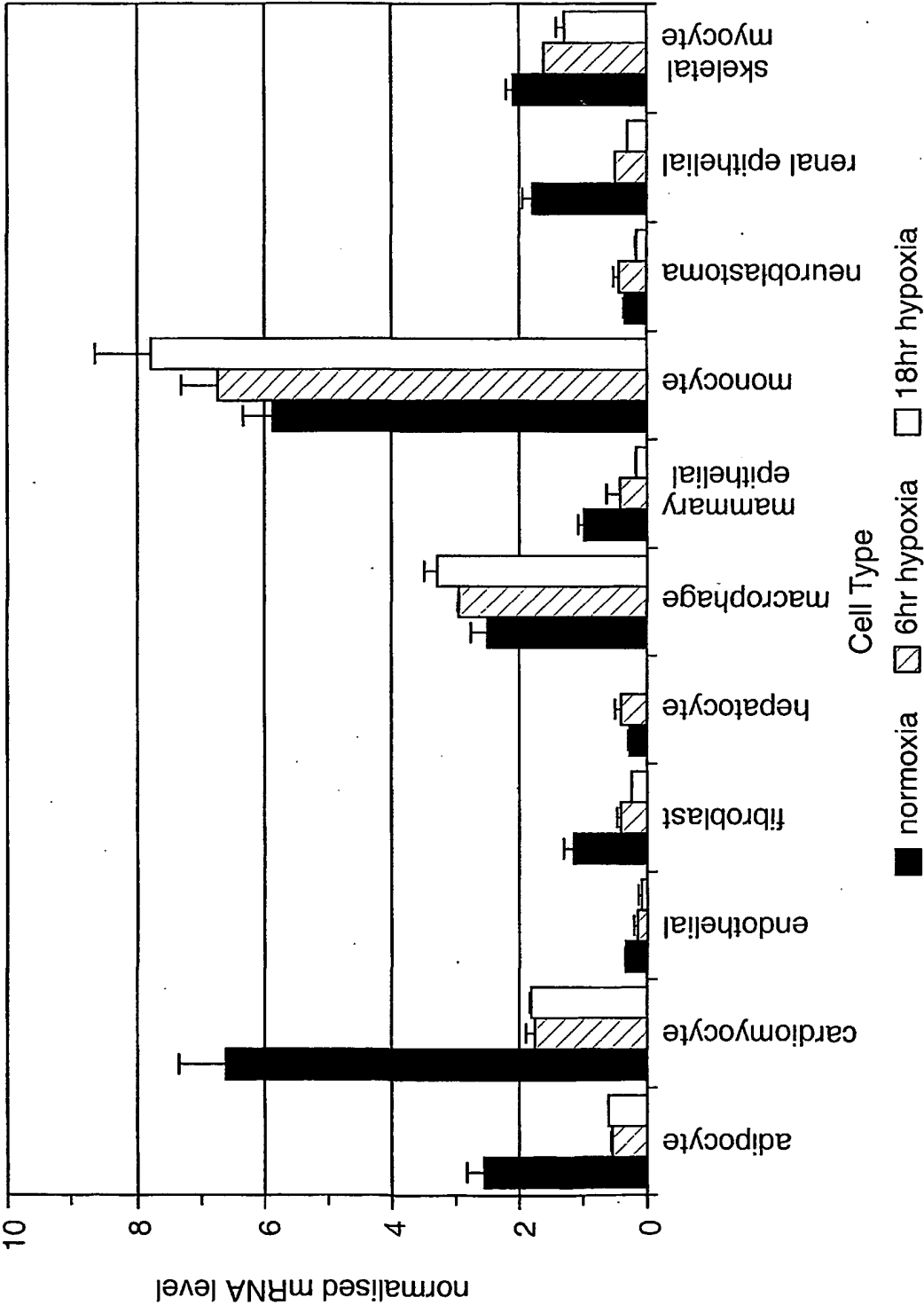
24/96

FIG. 19 p1F16/ SeqID:326/ CYP1B1



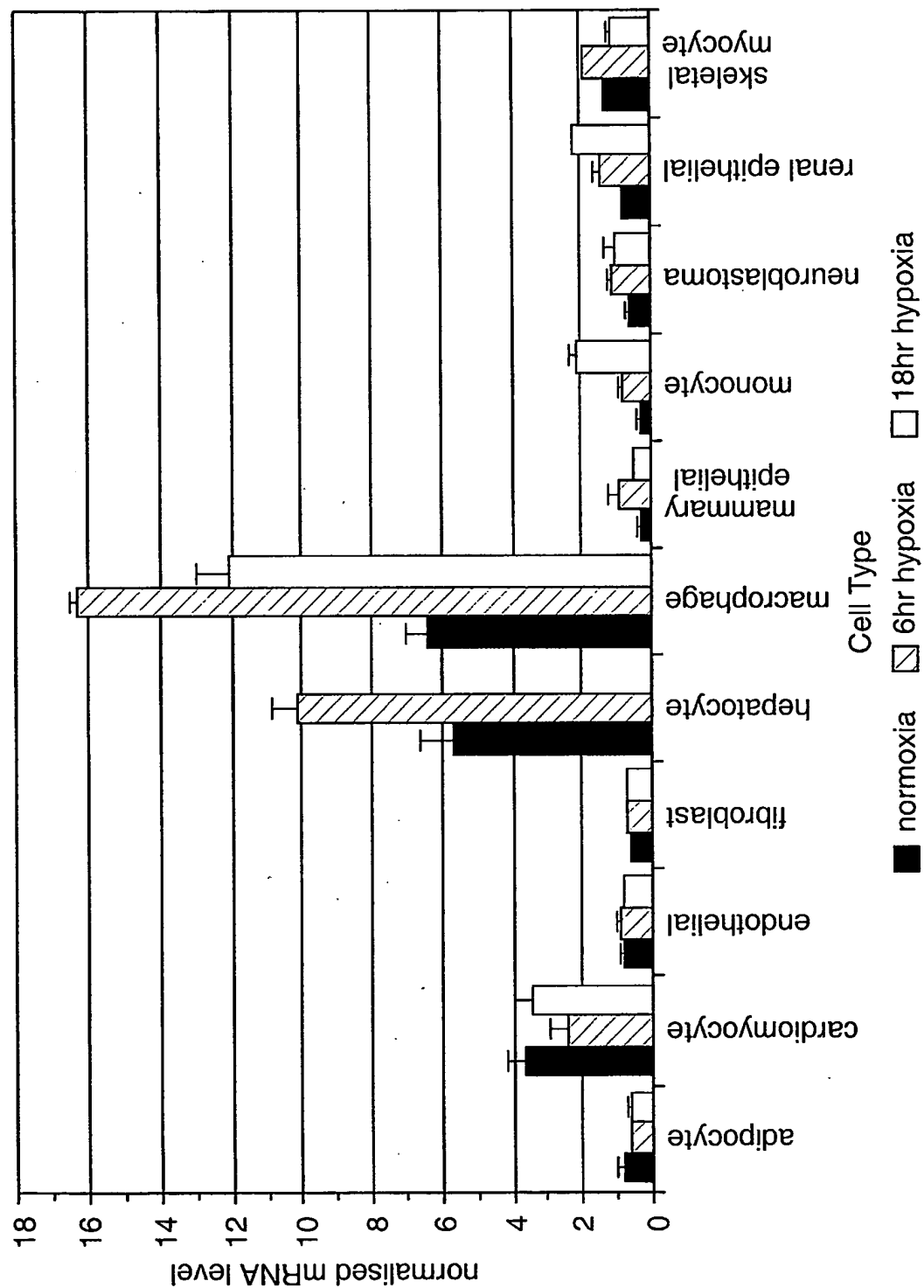
25/96

FIG. 20 p1E3/ SeqID:138/ CYP1B1



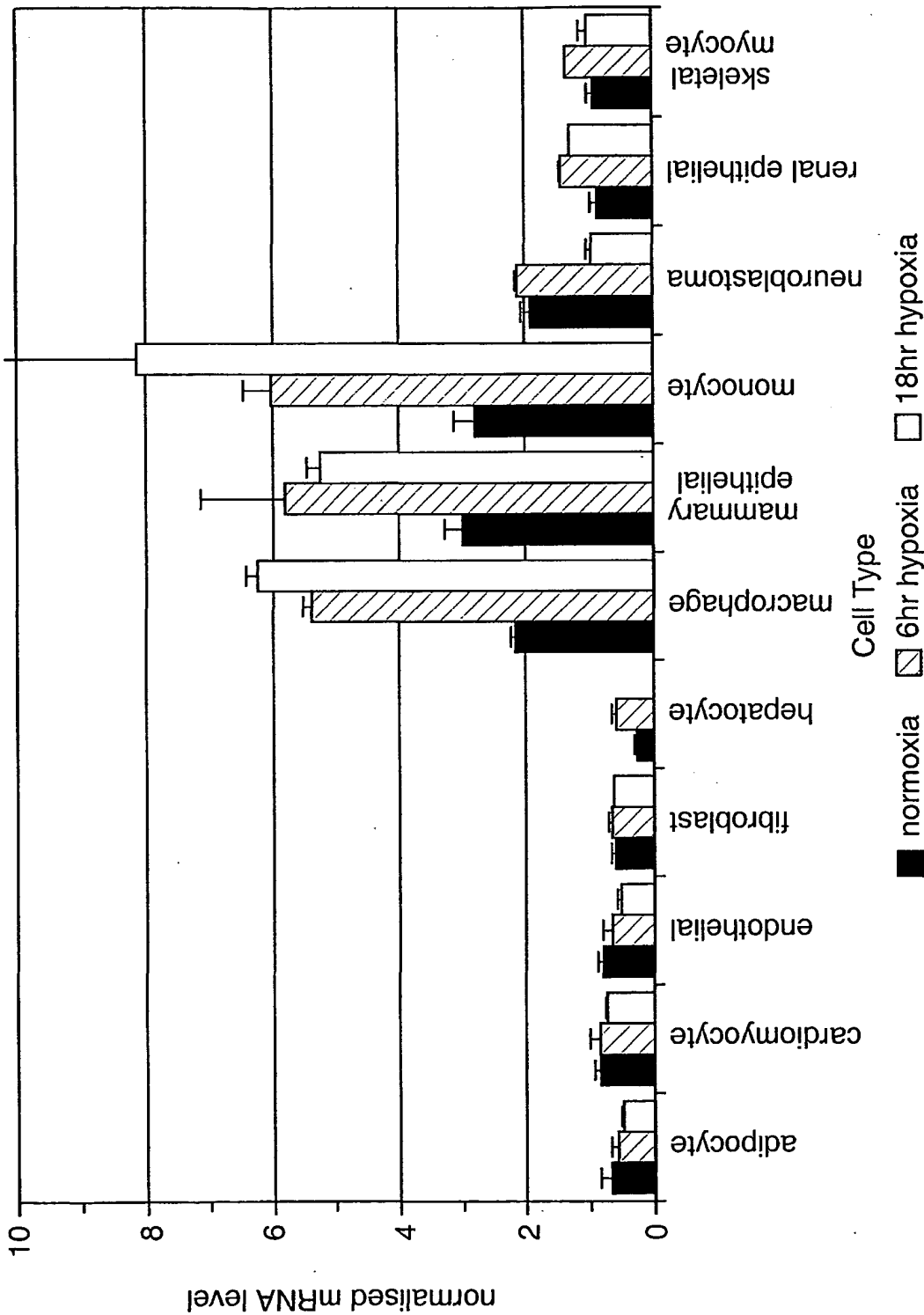
26/96

FIG. 21 p1H21/ SeqID:164/ Hypothetical protein FLJ13511(seq ID: 163/164)



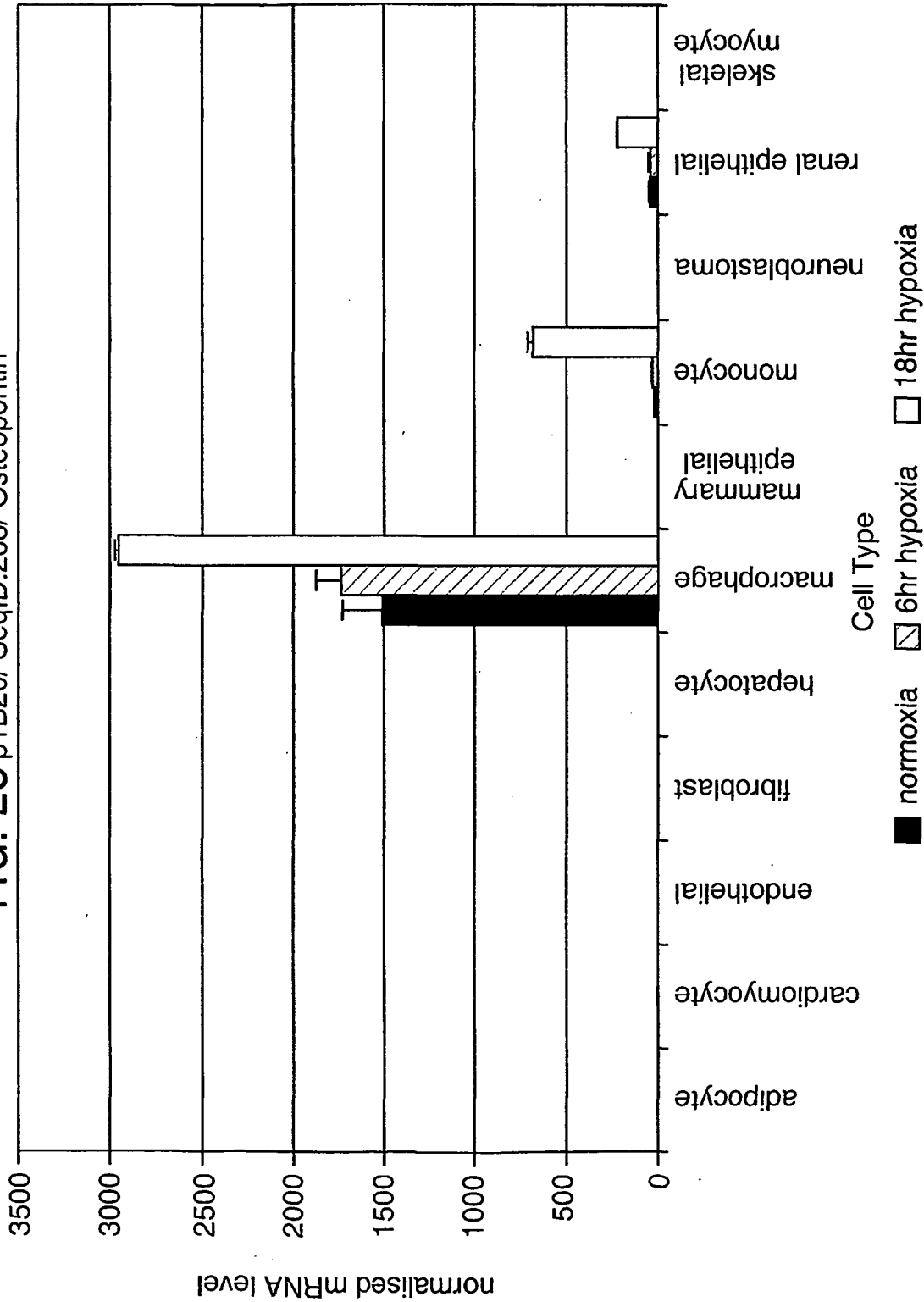
27/96

FIG. 22 p1F21/ SeqID:18/ Hematopoietic Zinc finger protein



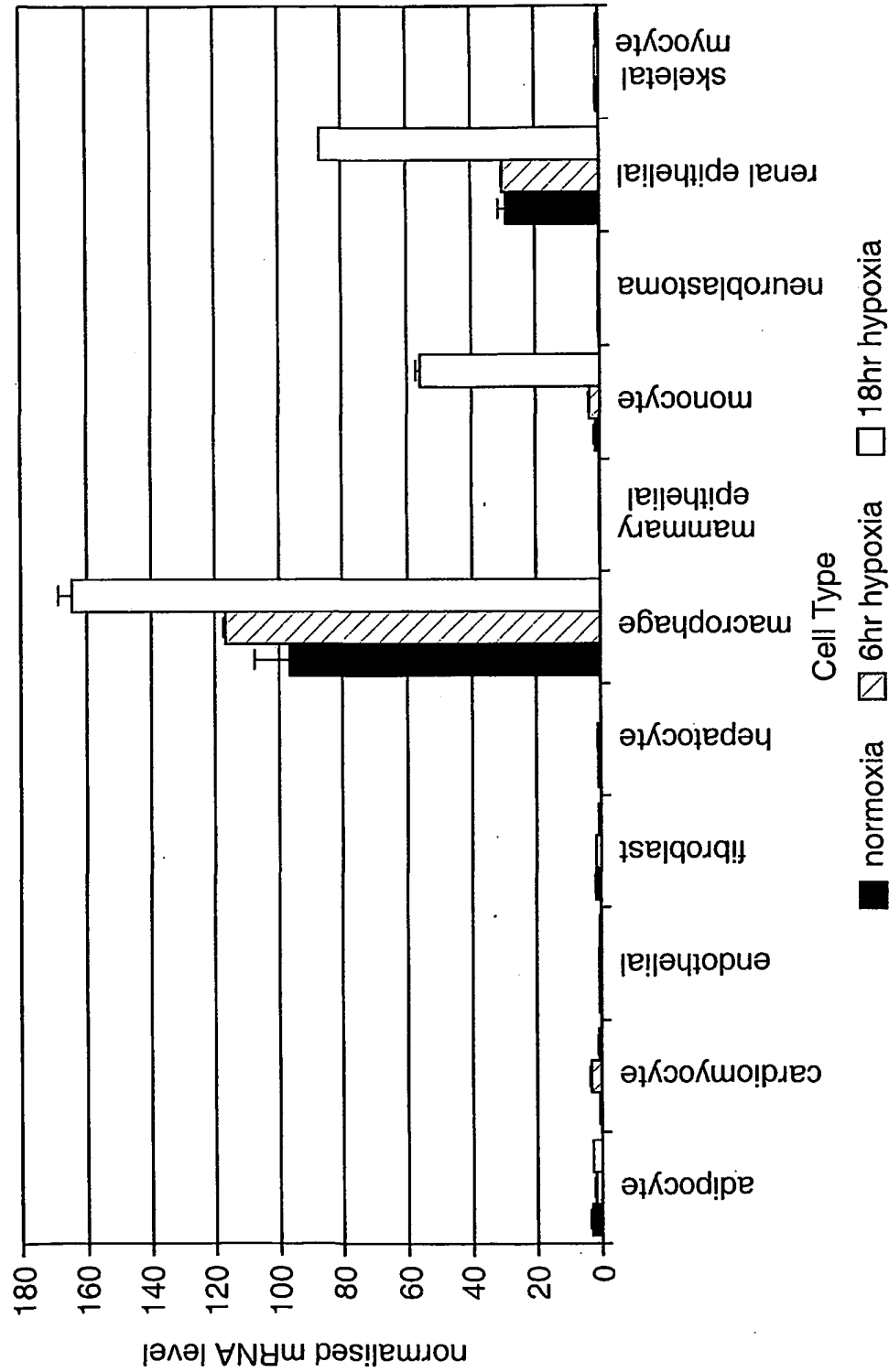
28/96

FIG. 23 p1B20/ SeqID:268/ Osteopontin



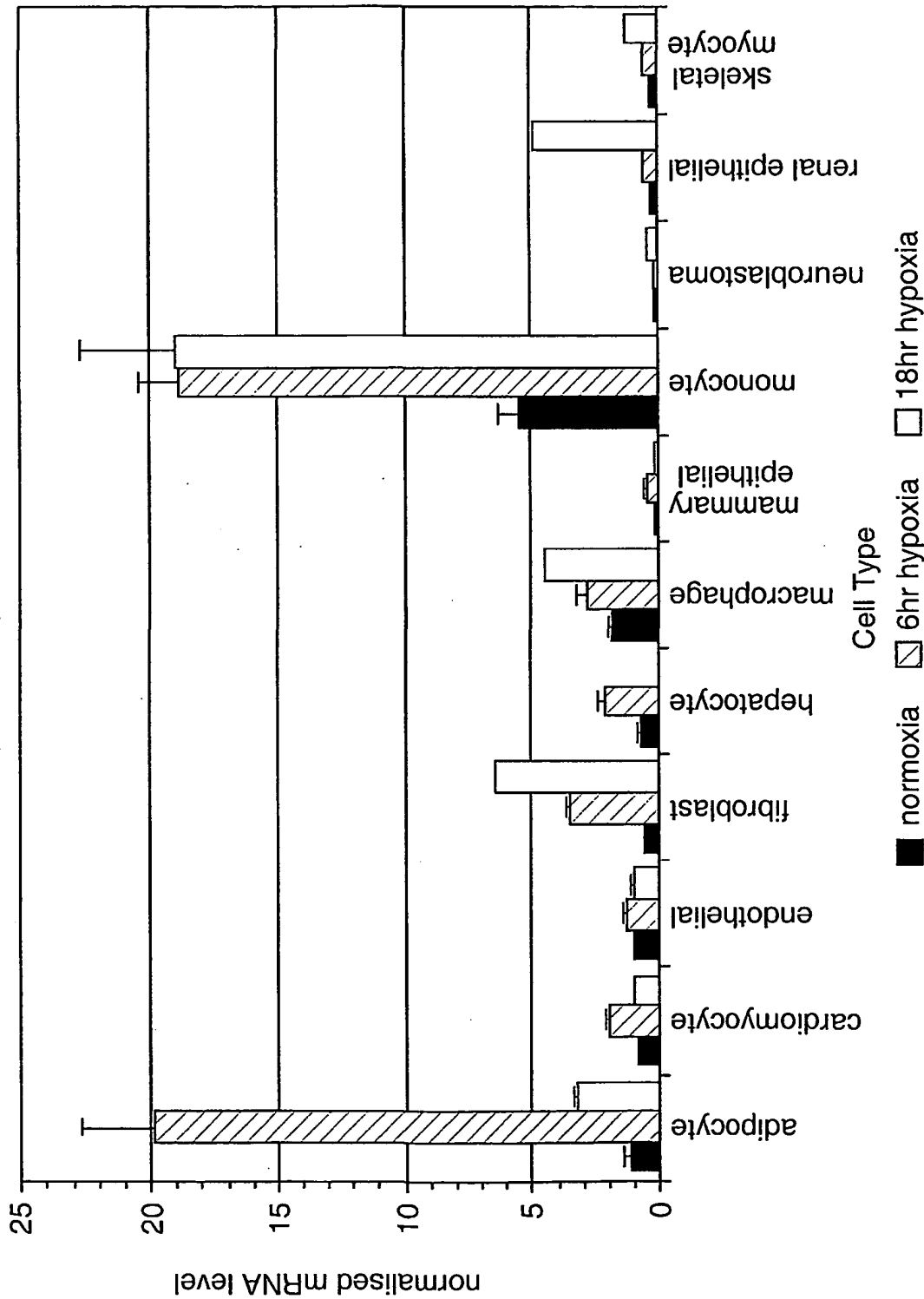
29/96

FIG. 24 p1B21/ SeqID:268/ Osteopontin



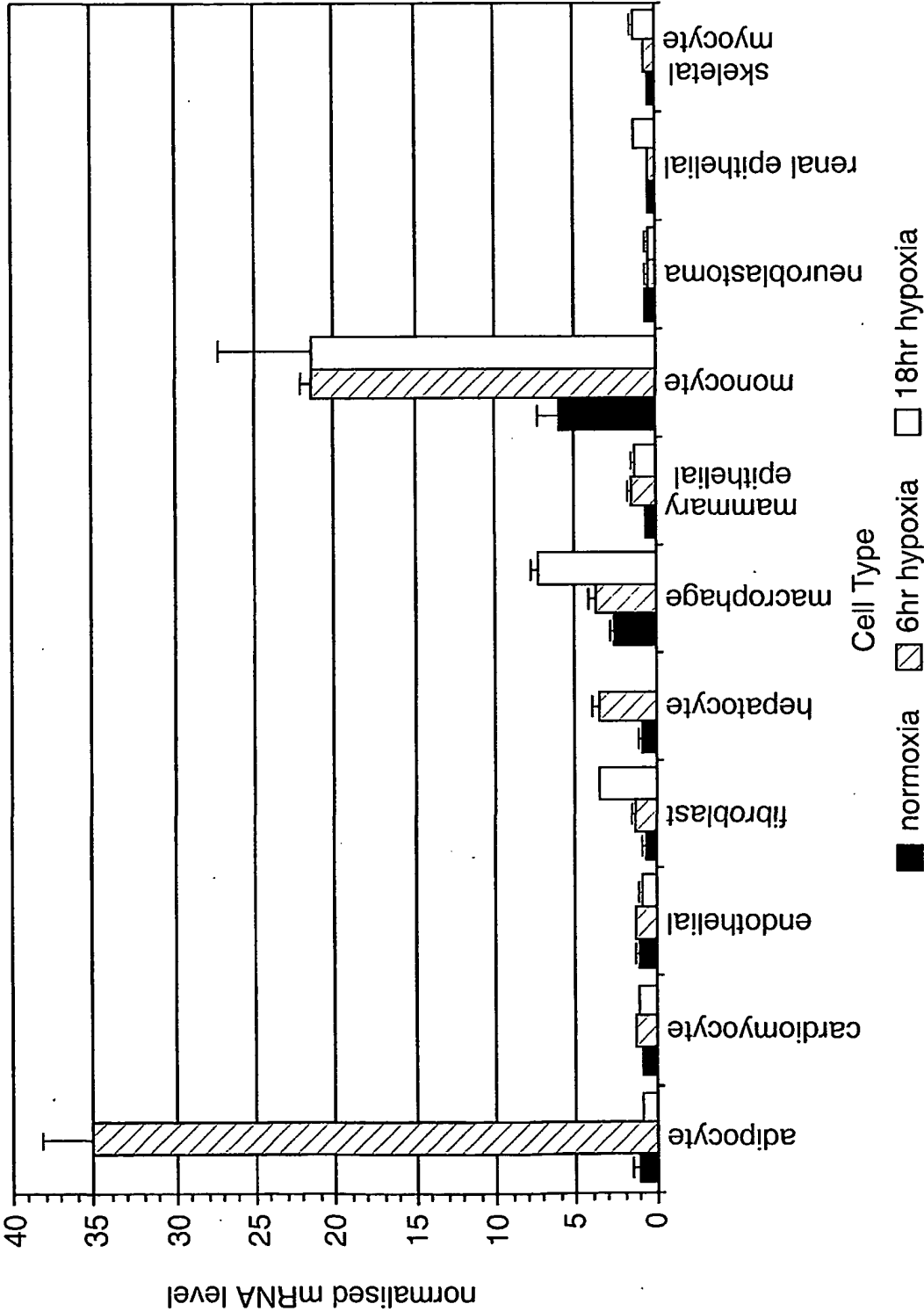
30/96

FIG. 25_{p1B9/ SeqID:314/ adipophilin}



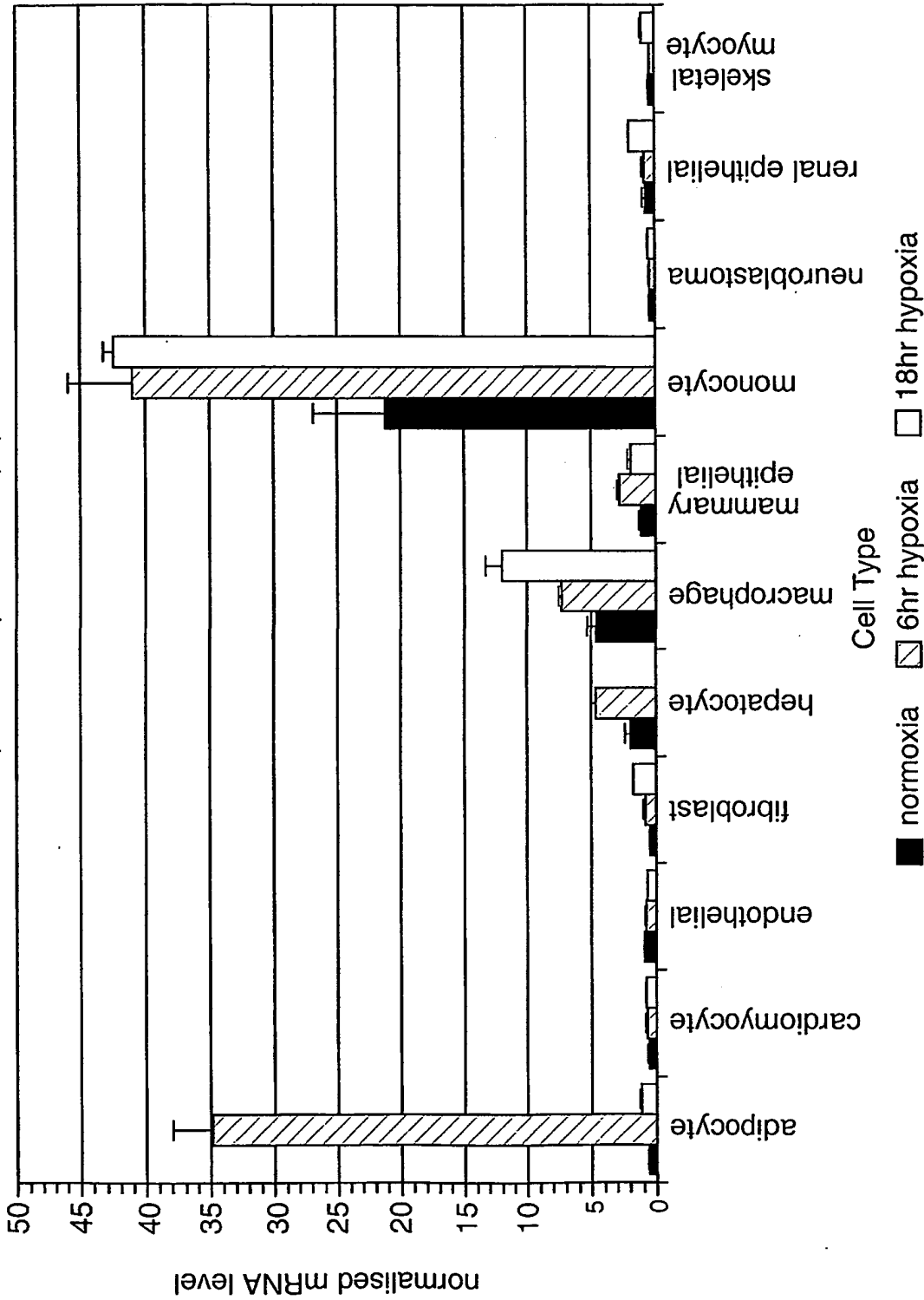
31/96

FIG. 26 p1B8/ SeqID:314/ adipophilin



32/96

FIG. 27 p1B7/ SeqID:314/ adipophilin



33/96

FIG. 28 p1B6/ SeqID:314/ adipophilin

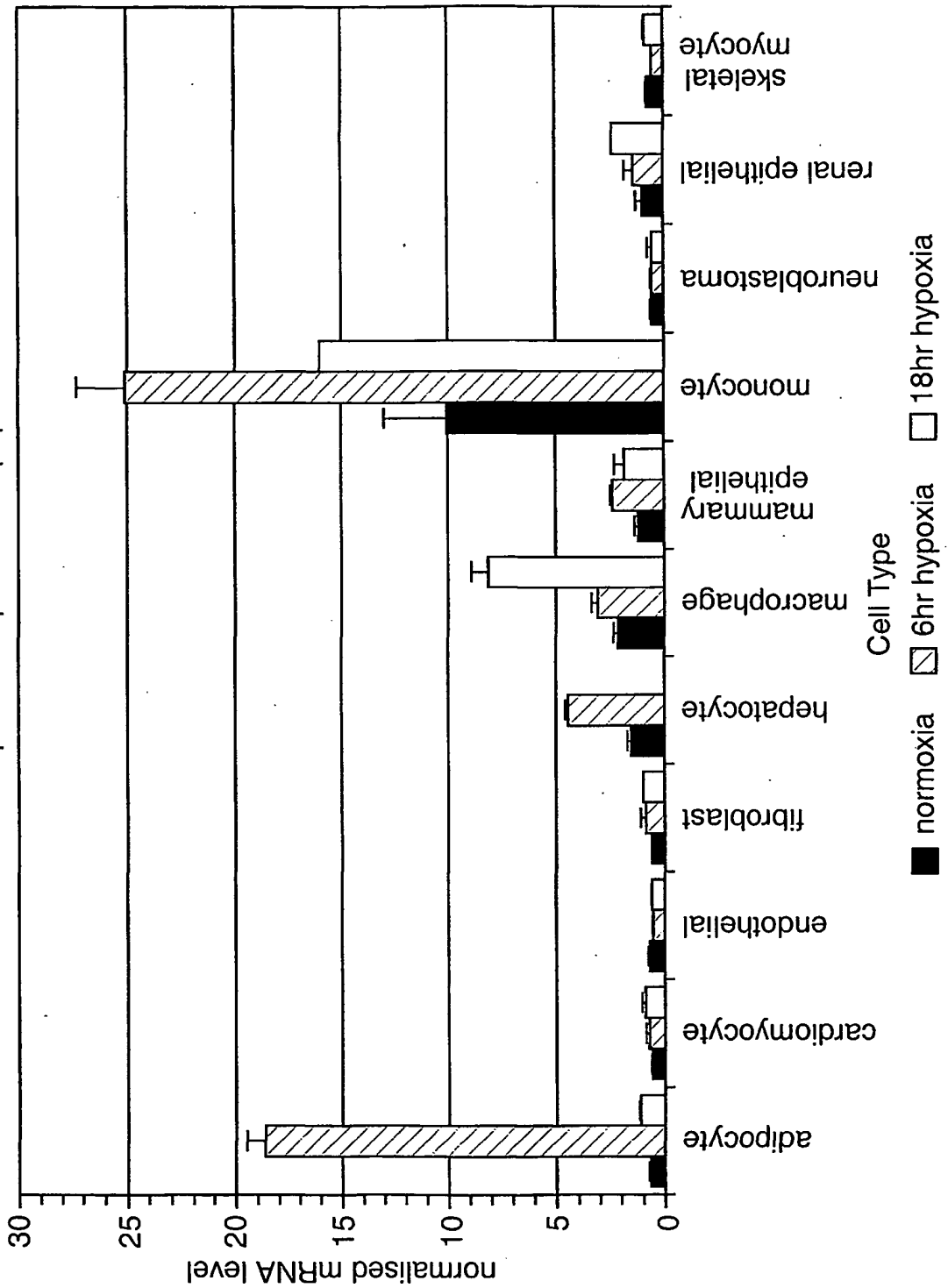
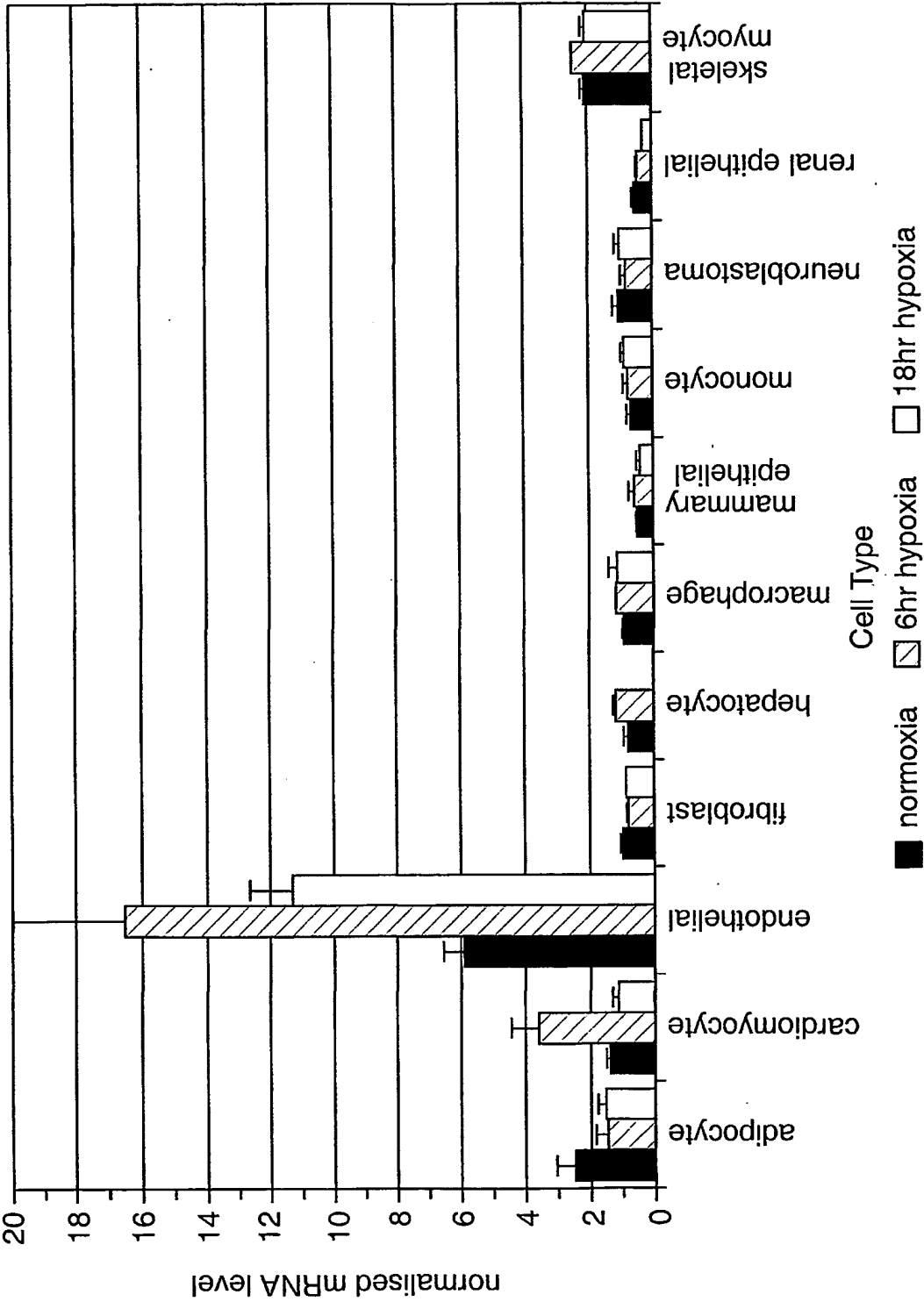


FIG. 29 p1H5/ SeqID:206/ Hypothetical protein FLJ22690



35/96

FIG. 30 p1E16/ SeqID:66/ cDNA DKFZp586E1624

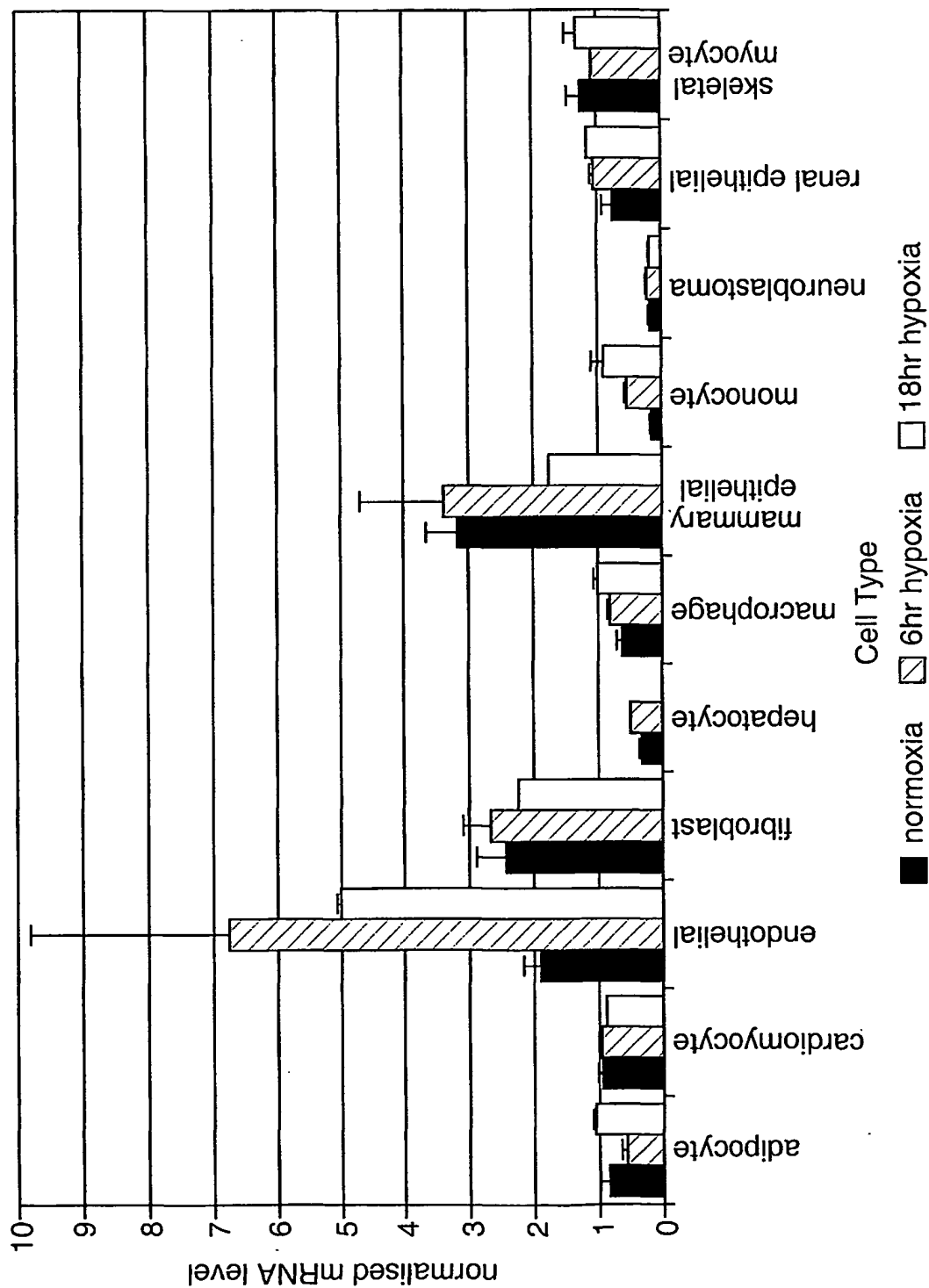
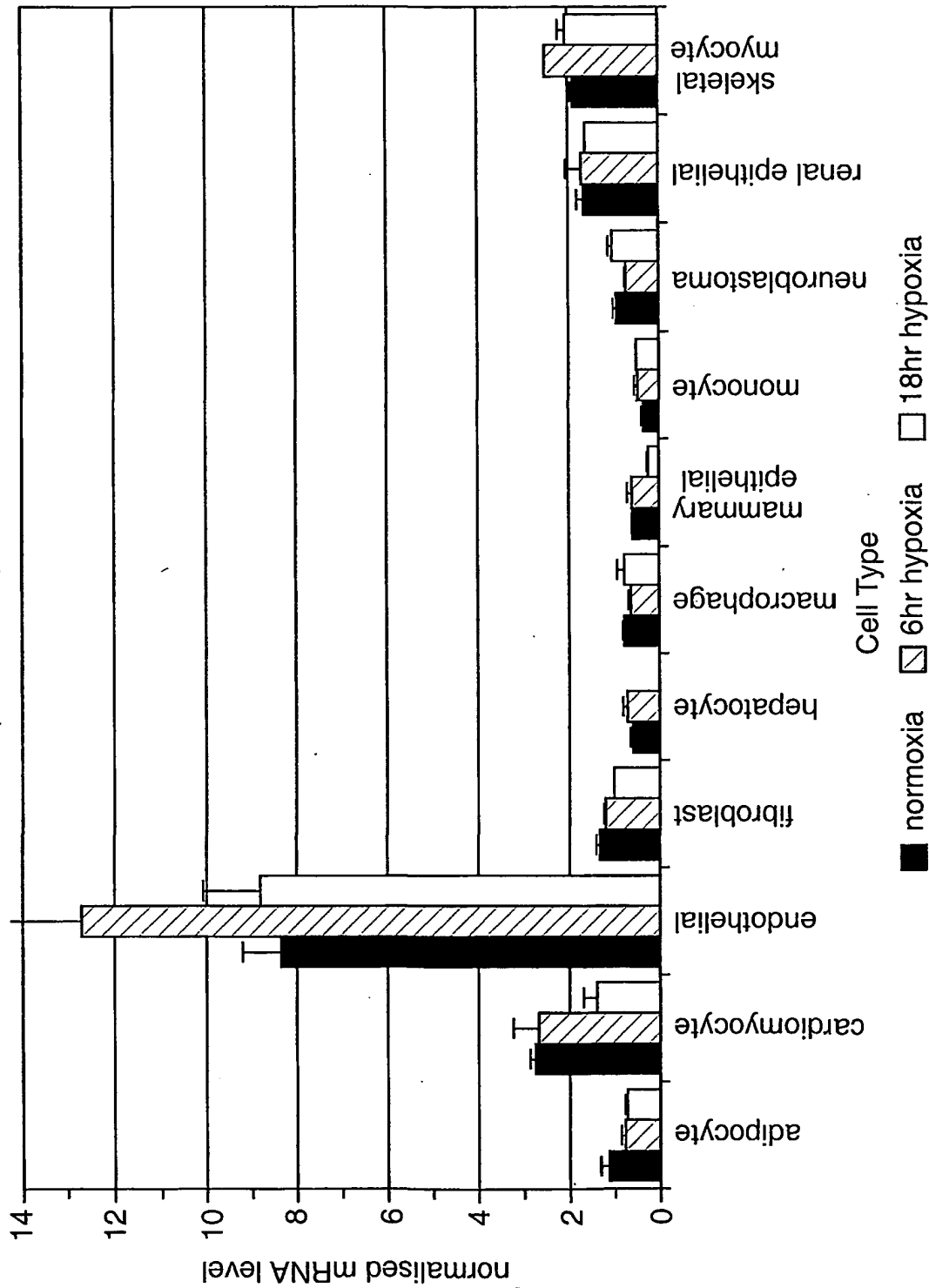


FIG. 31 p1G22/ SeqID:198/ EST



37/96

FIG. 32a p1E6/ SeqID:86/ EGL nine (C.elegans) homolog 3

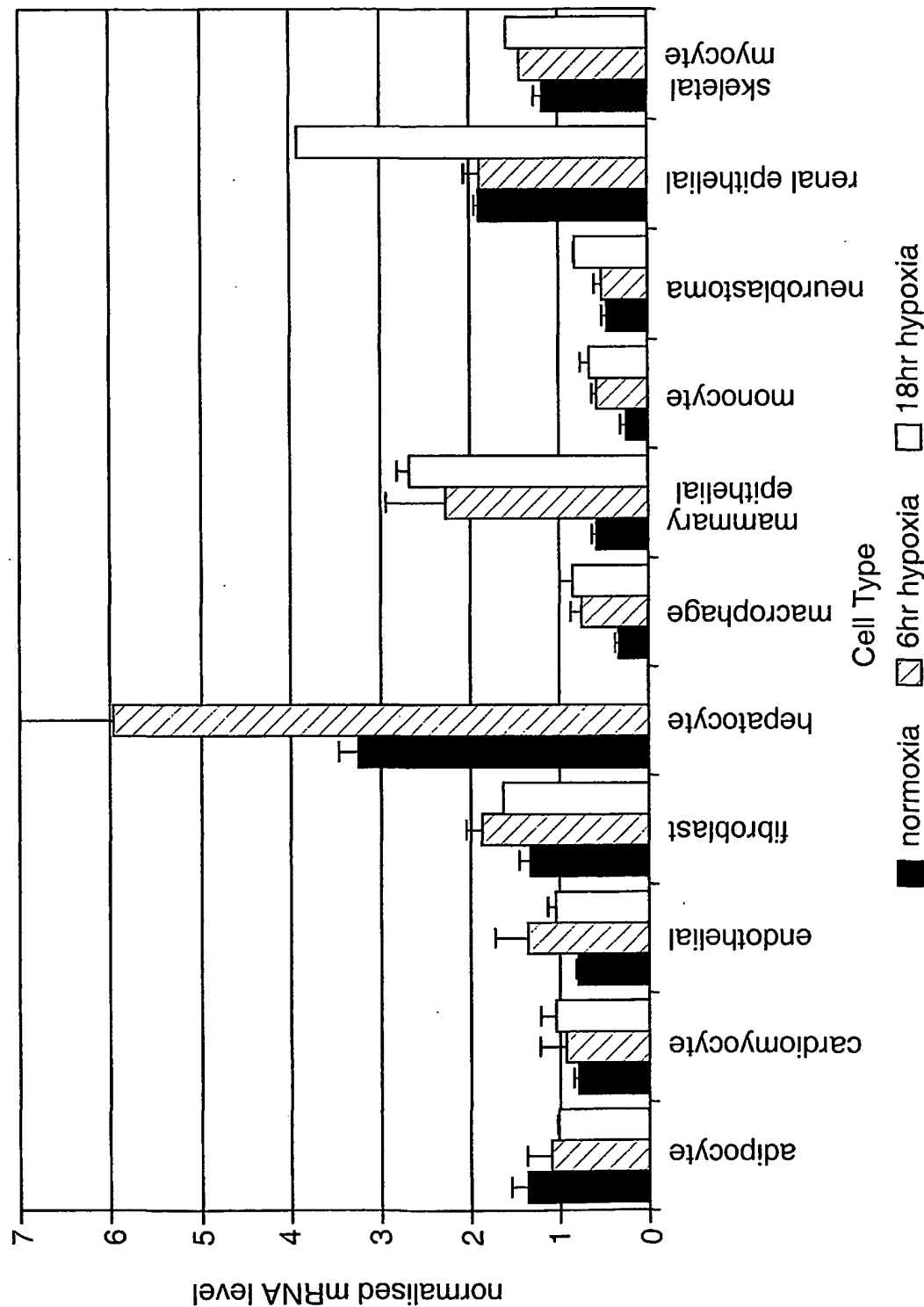


FIG. 32c

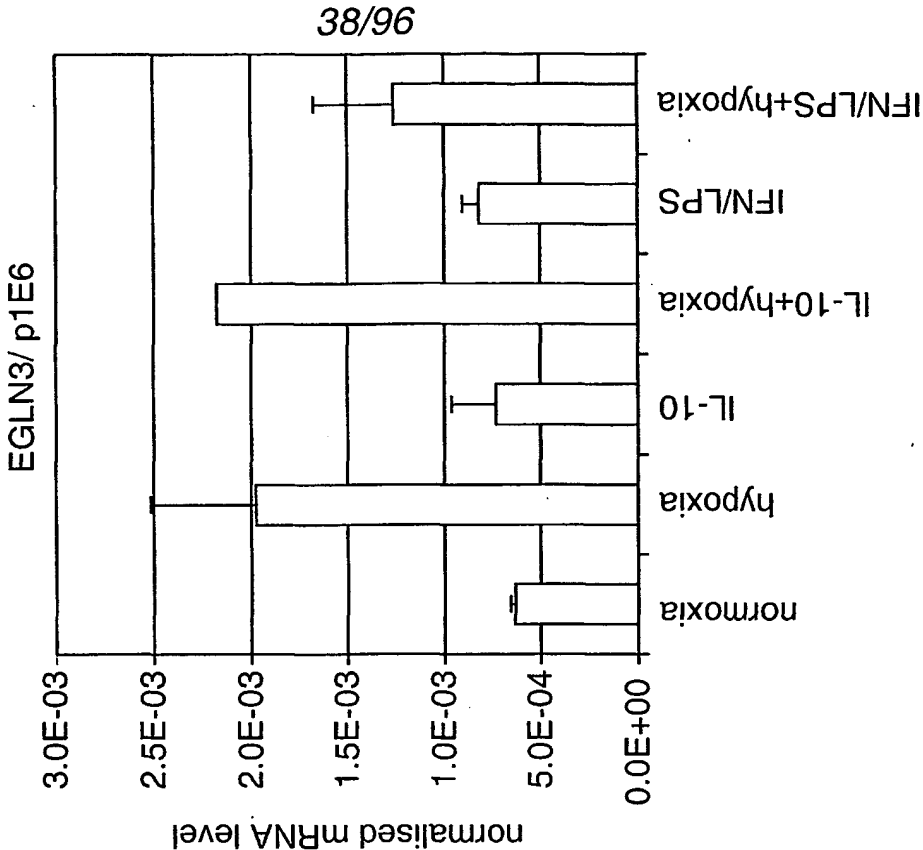
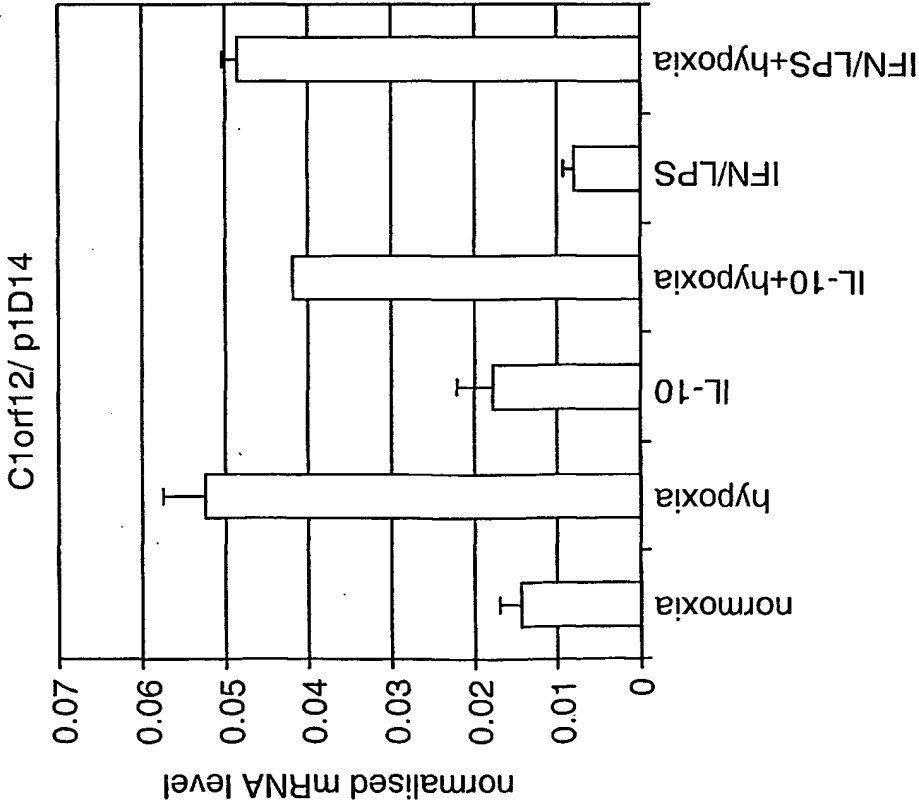
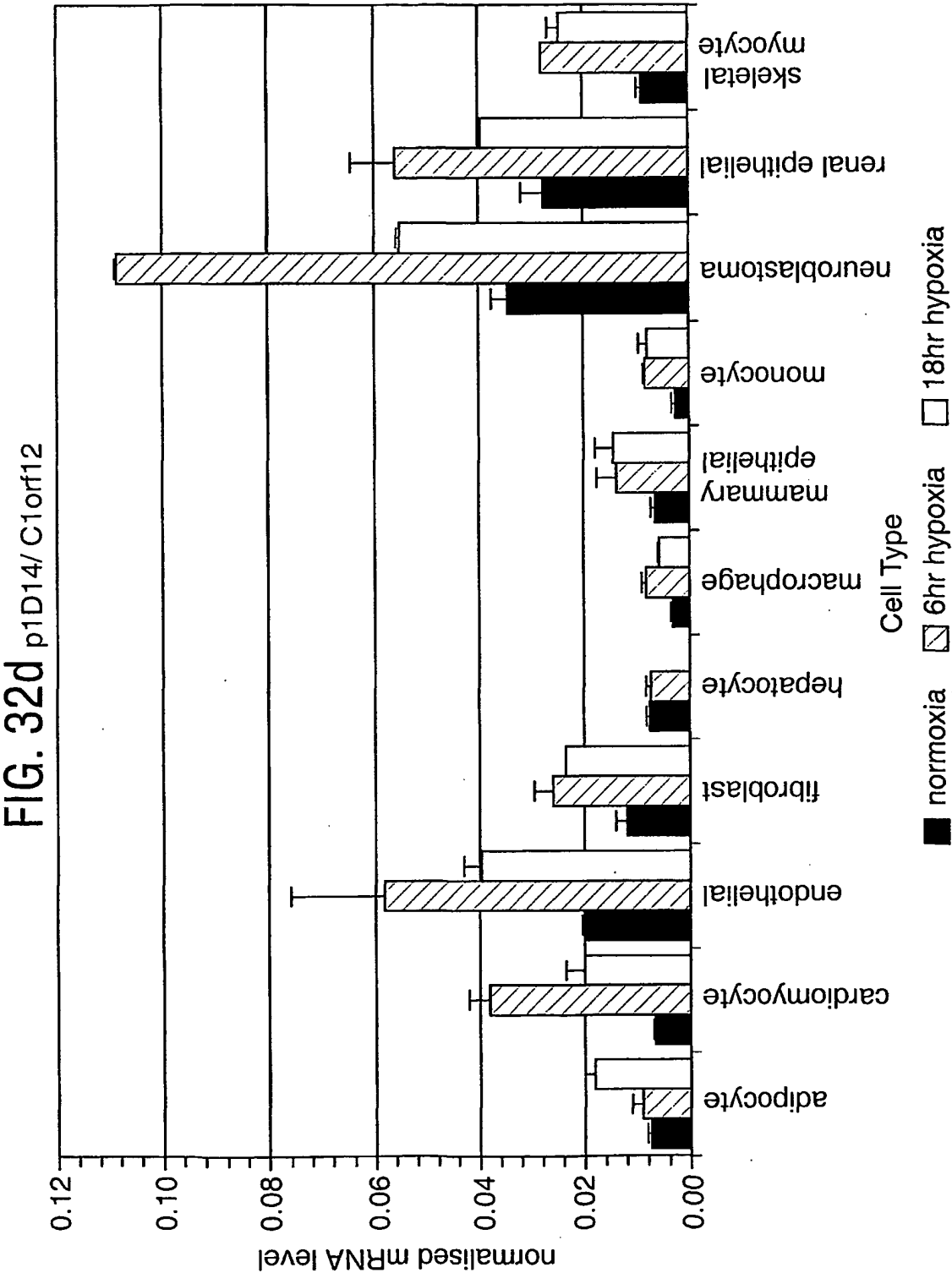


FIG. 32b



39/96



40/96

FIG. 32e

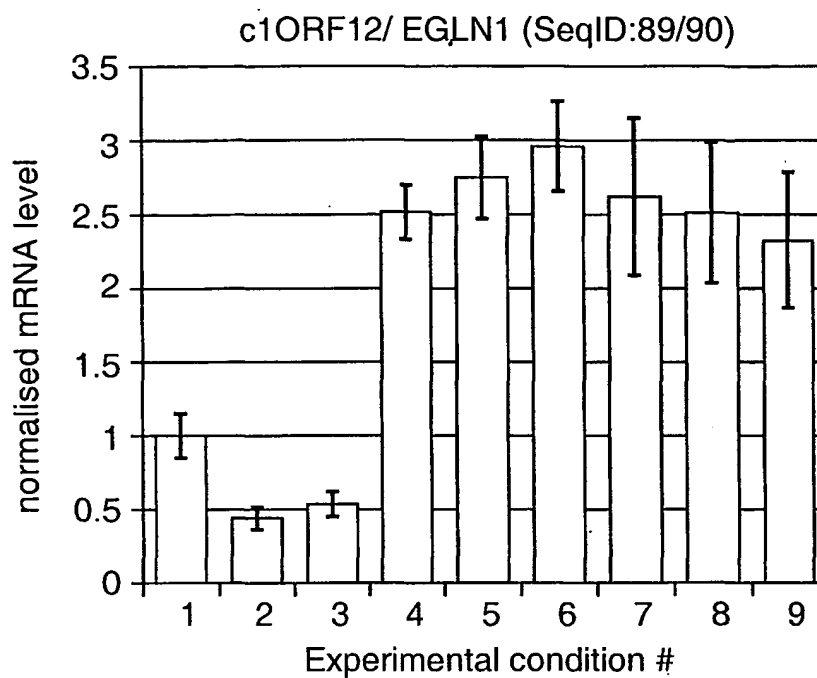
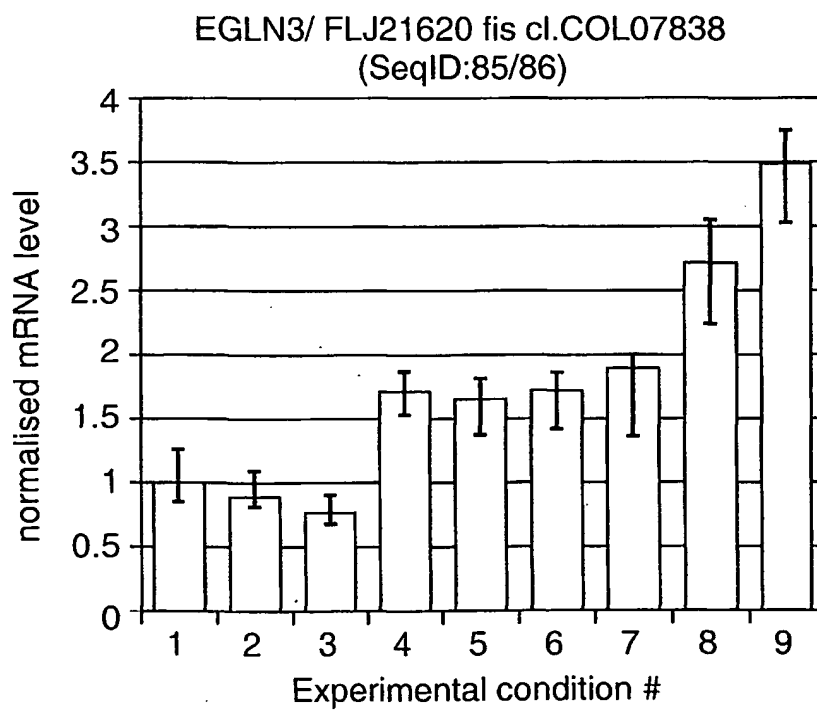
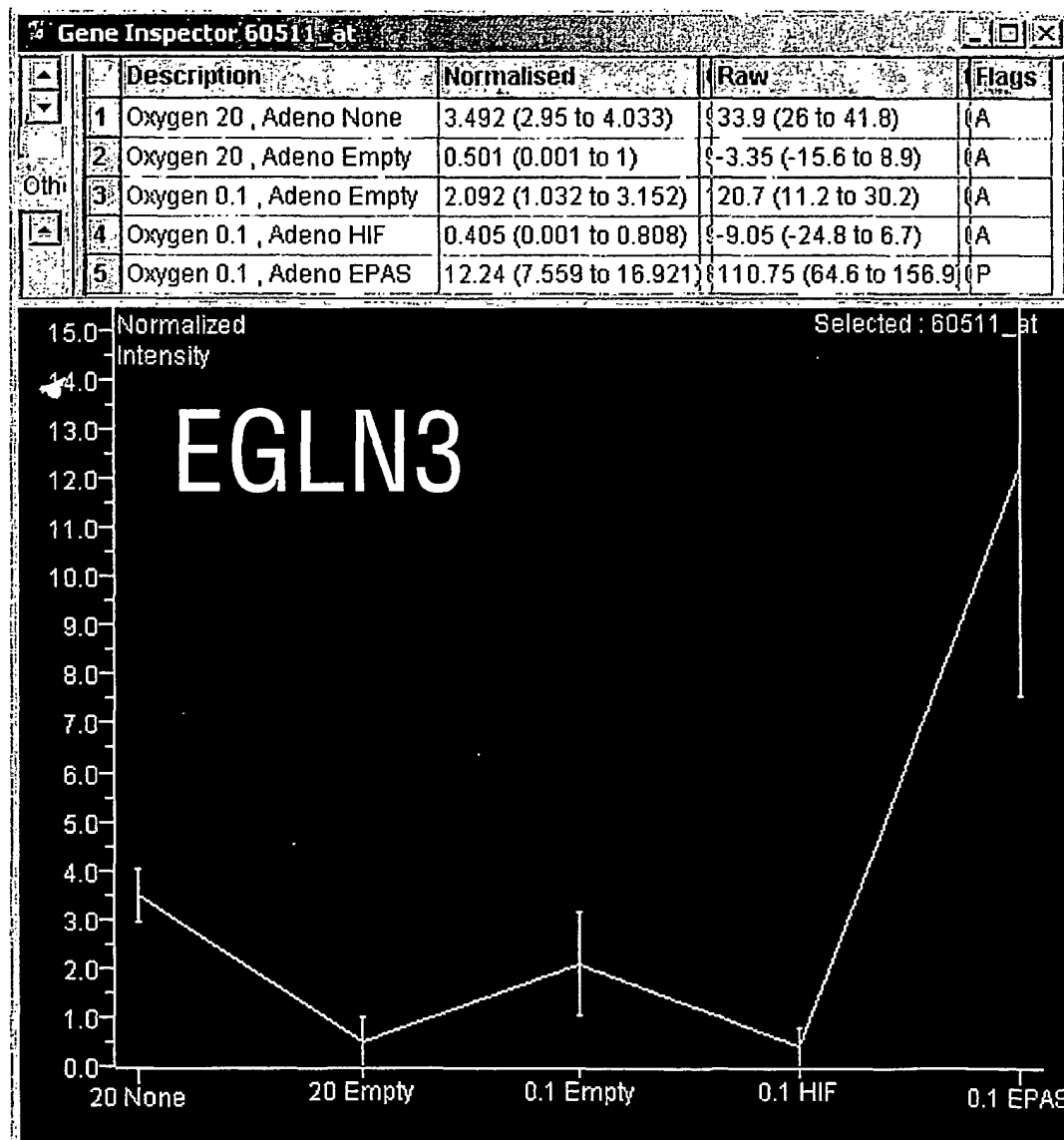


FIG. 32f



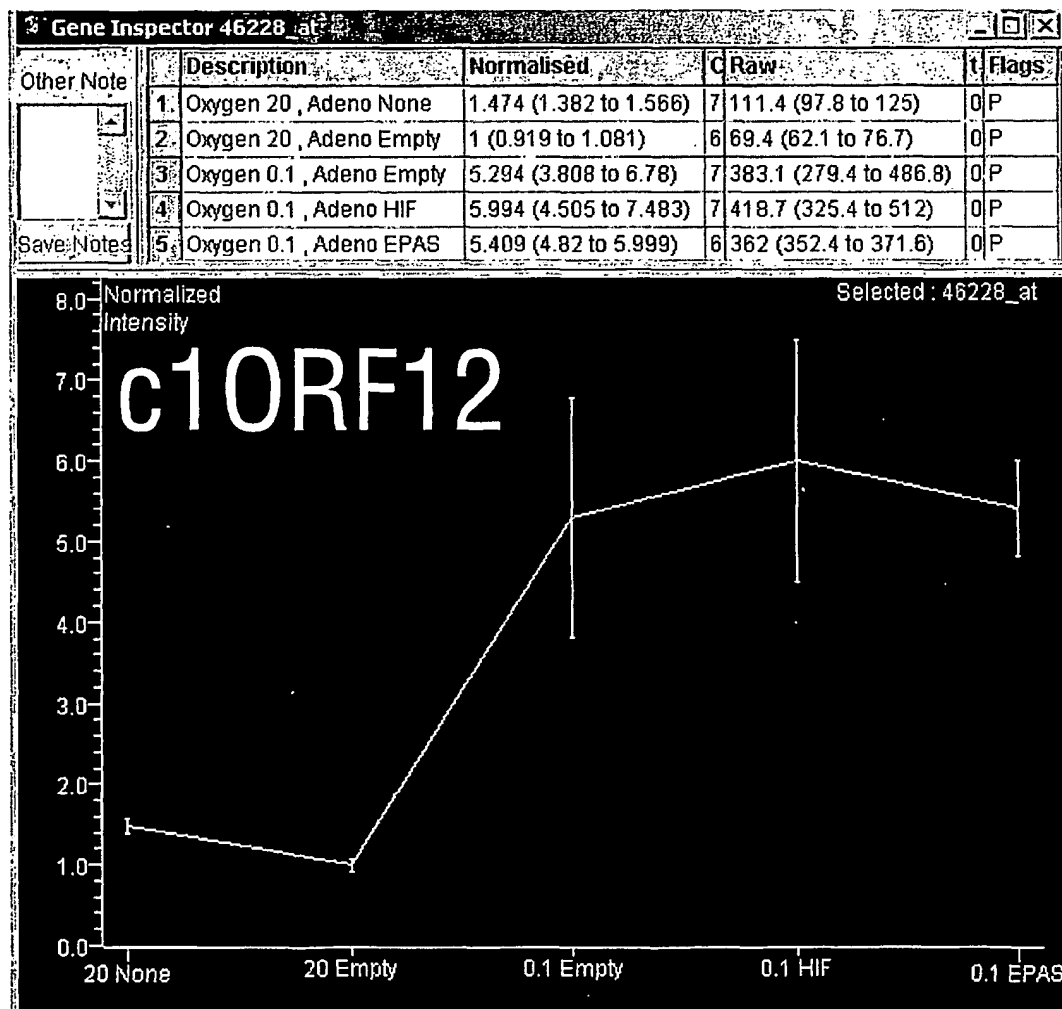
41/96

FIG. 32g



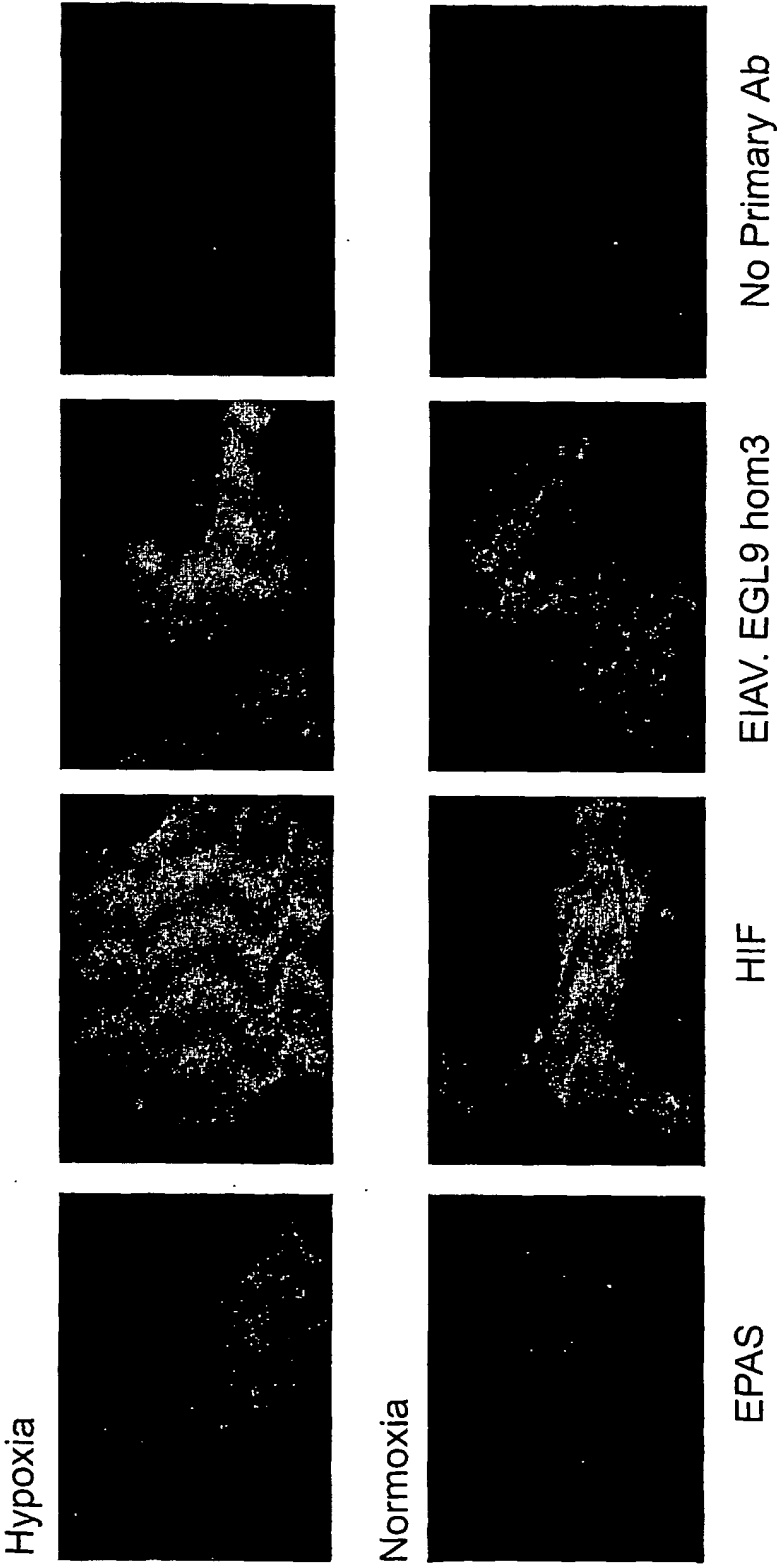
42/96

FIG. 32h

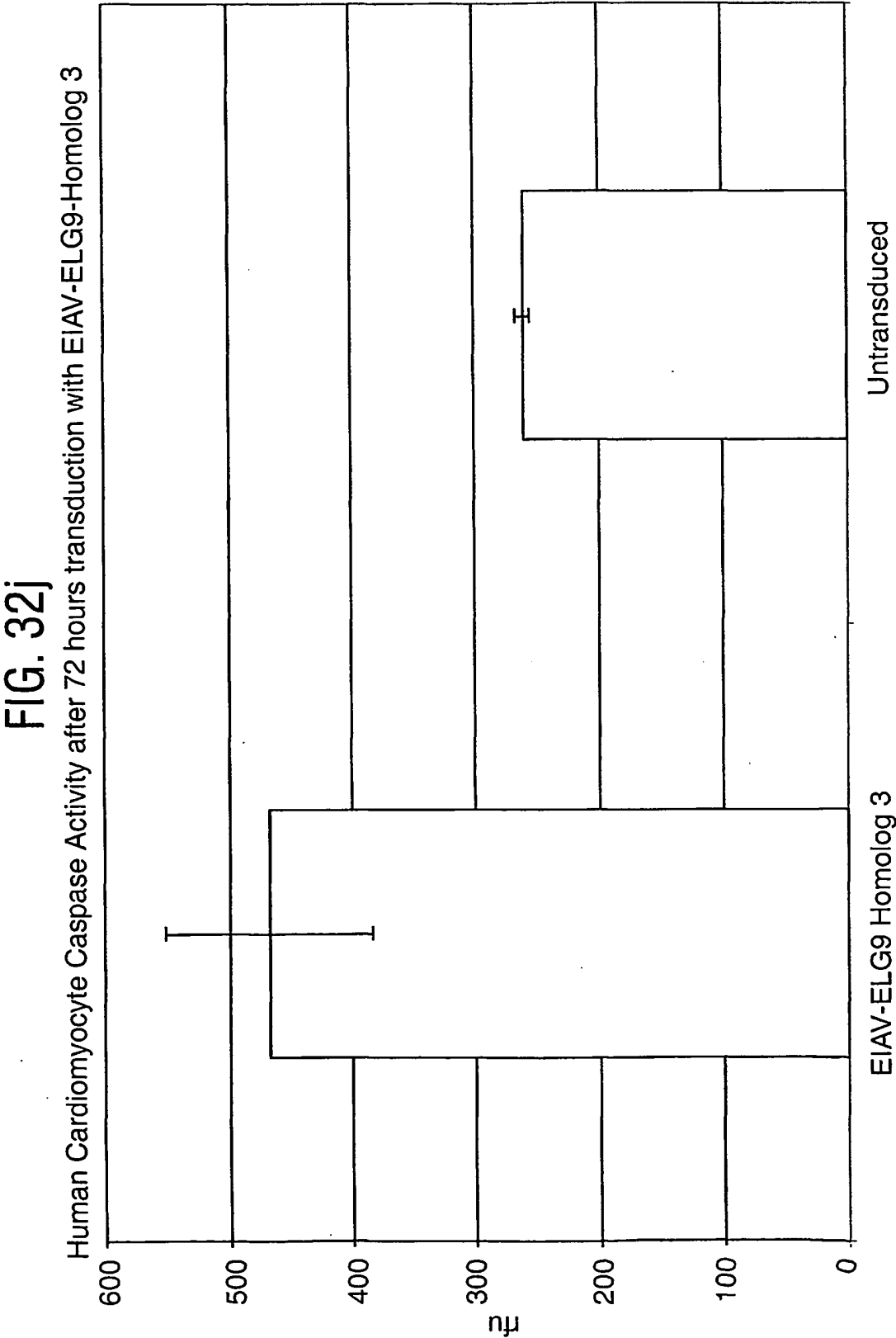


43/96

FIG. 32i
Flag immunocytochemistry in HEK293T cells



44/96



45/96

FIG. 33 p1E7/ SeqID:84/ Novel Metallothionein

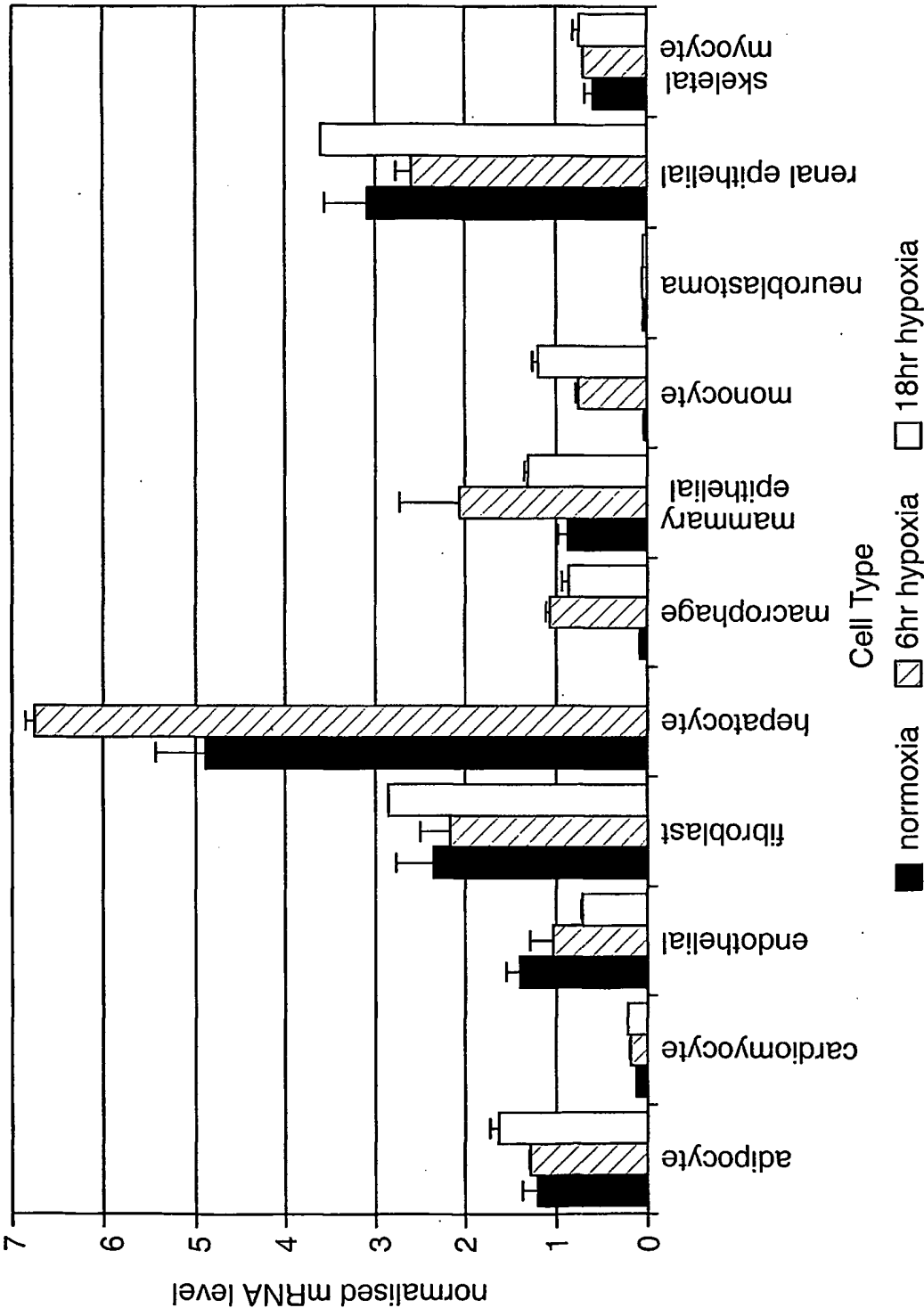
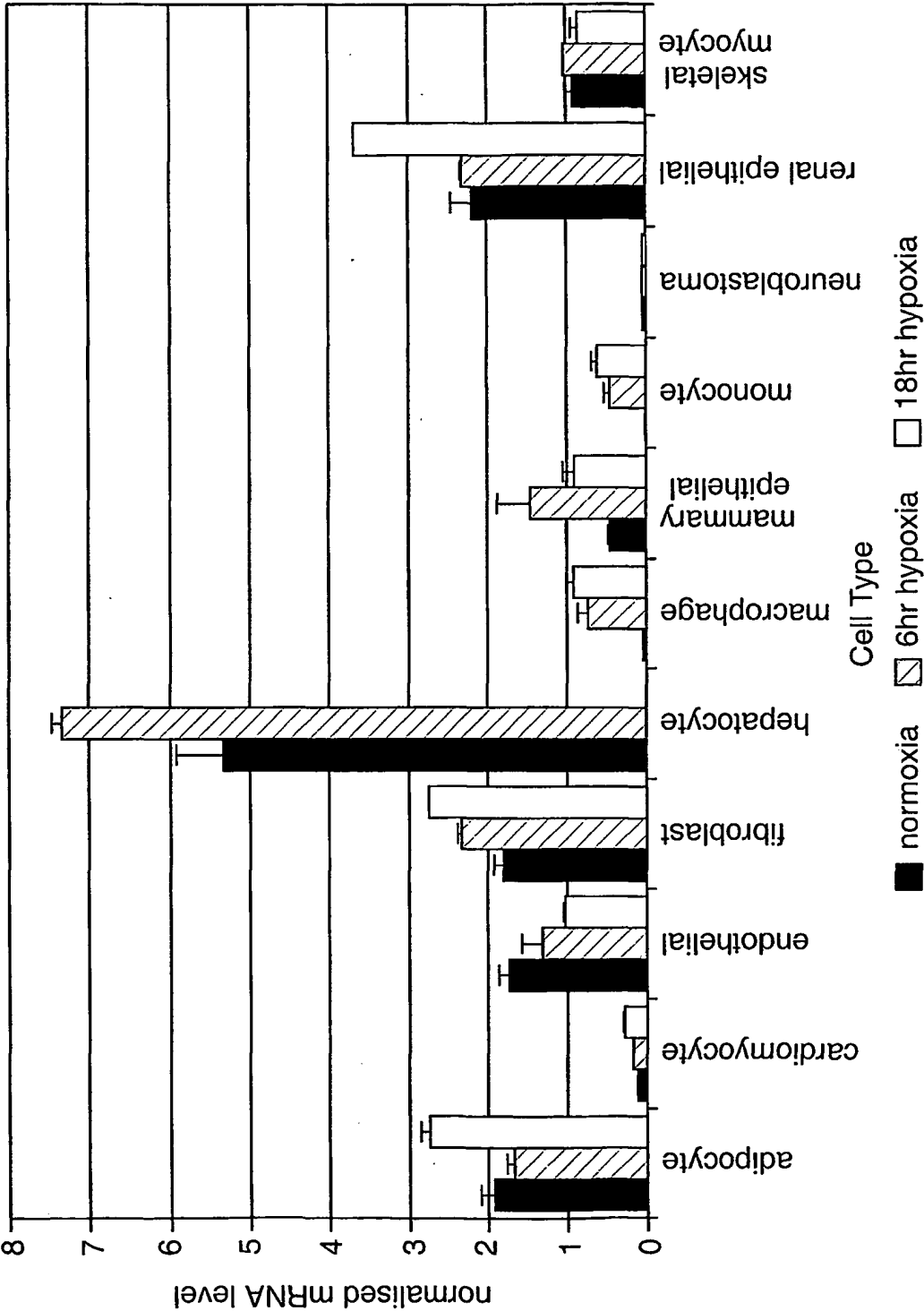
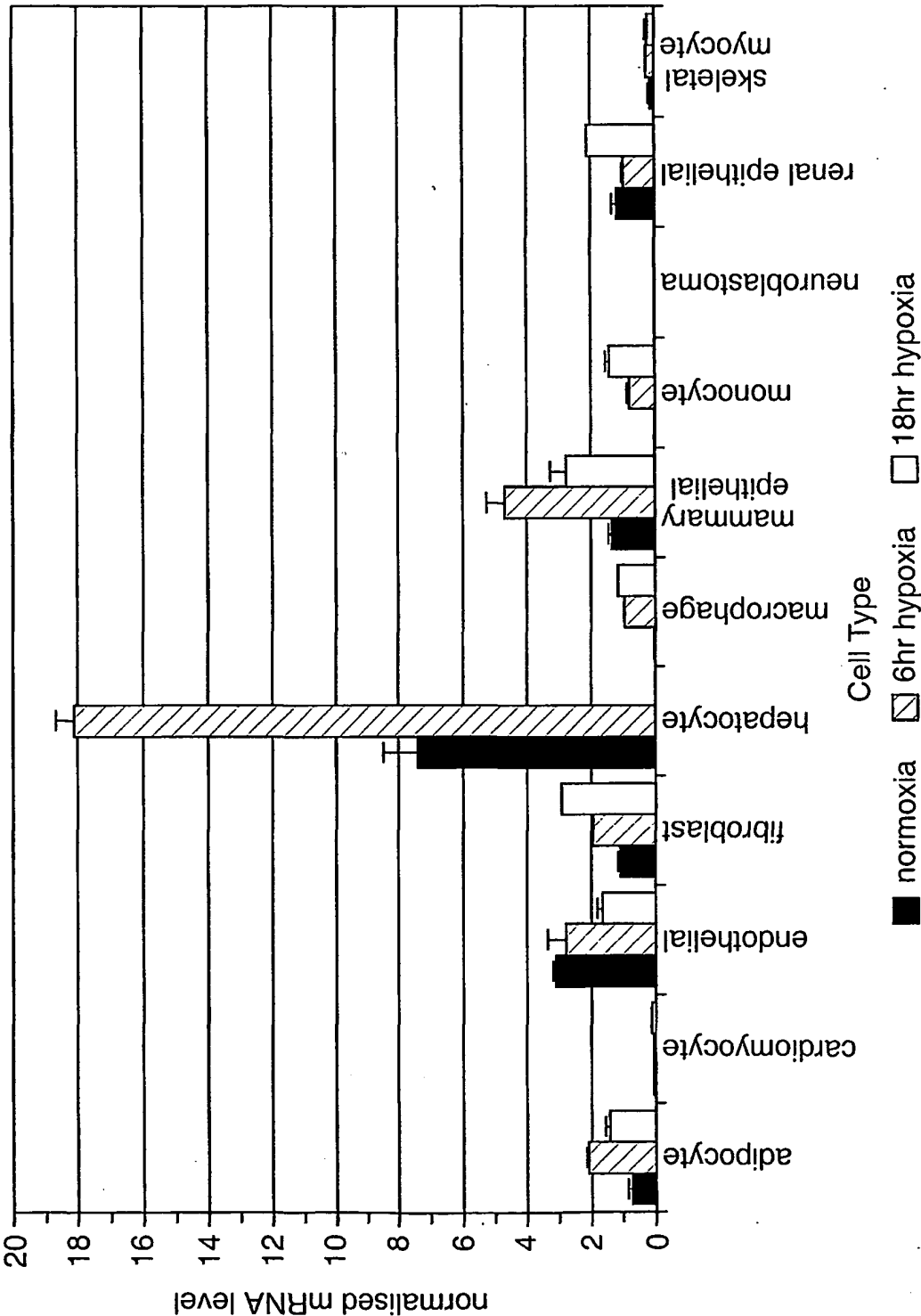


FIG. 34 p1F6/ SeqID:338/ Hypothetical protein hqp0376



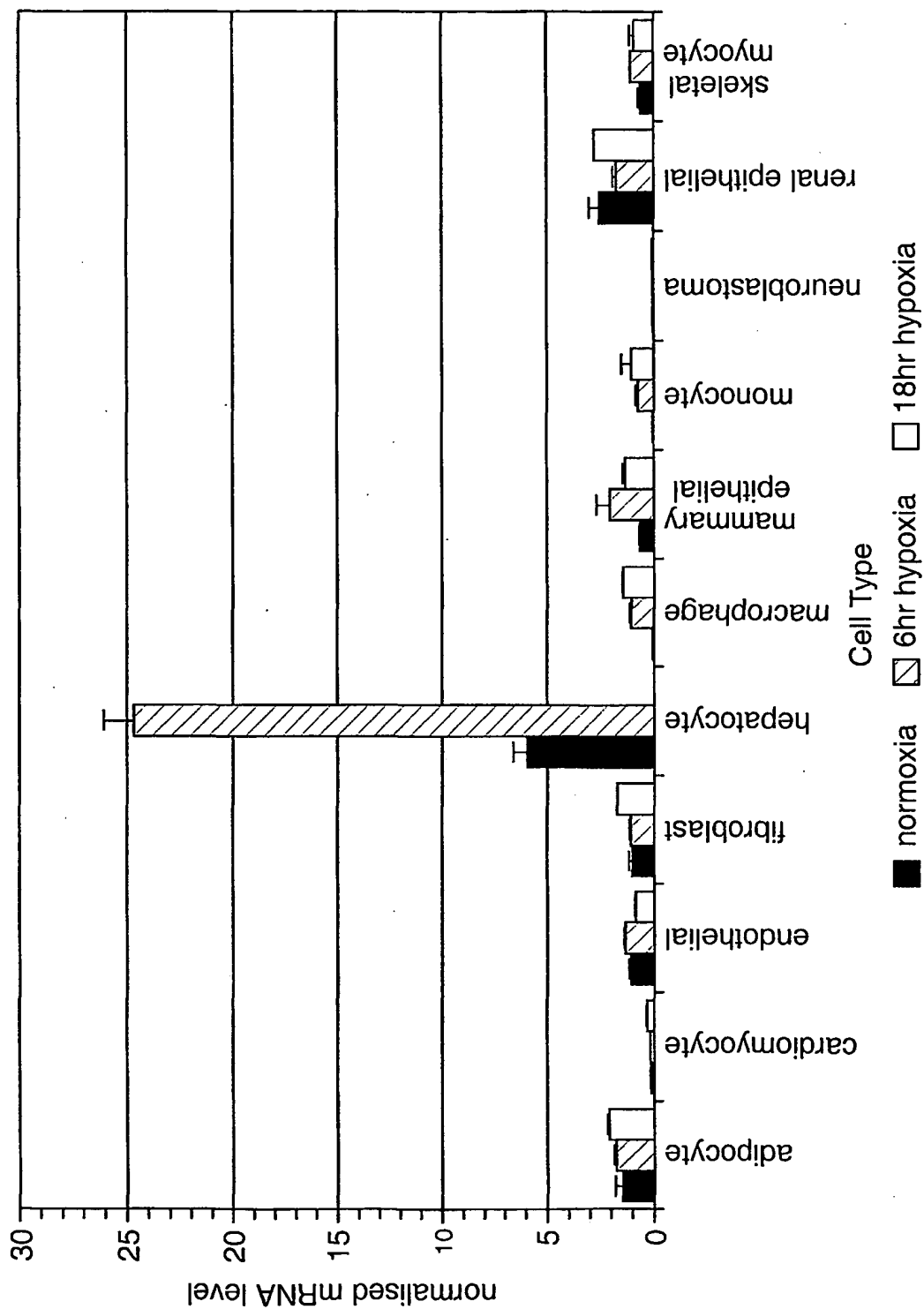
47/96

FIG. 35 p1A23/ SeqID:266/ Metallothionein 2A



48/96

FIG. 36 p1B1/ SeqID:244/ Metallothionein 1G



49/96

FIG. 37 p1A24/ SeqID:240/ Metallothionein 1H

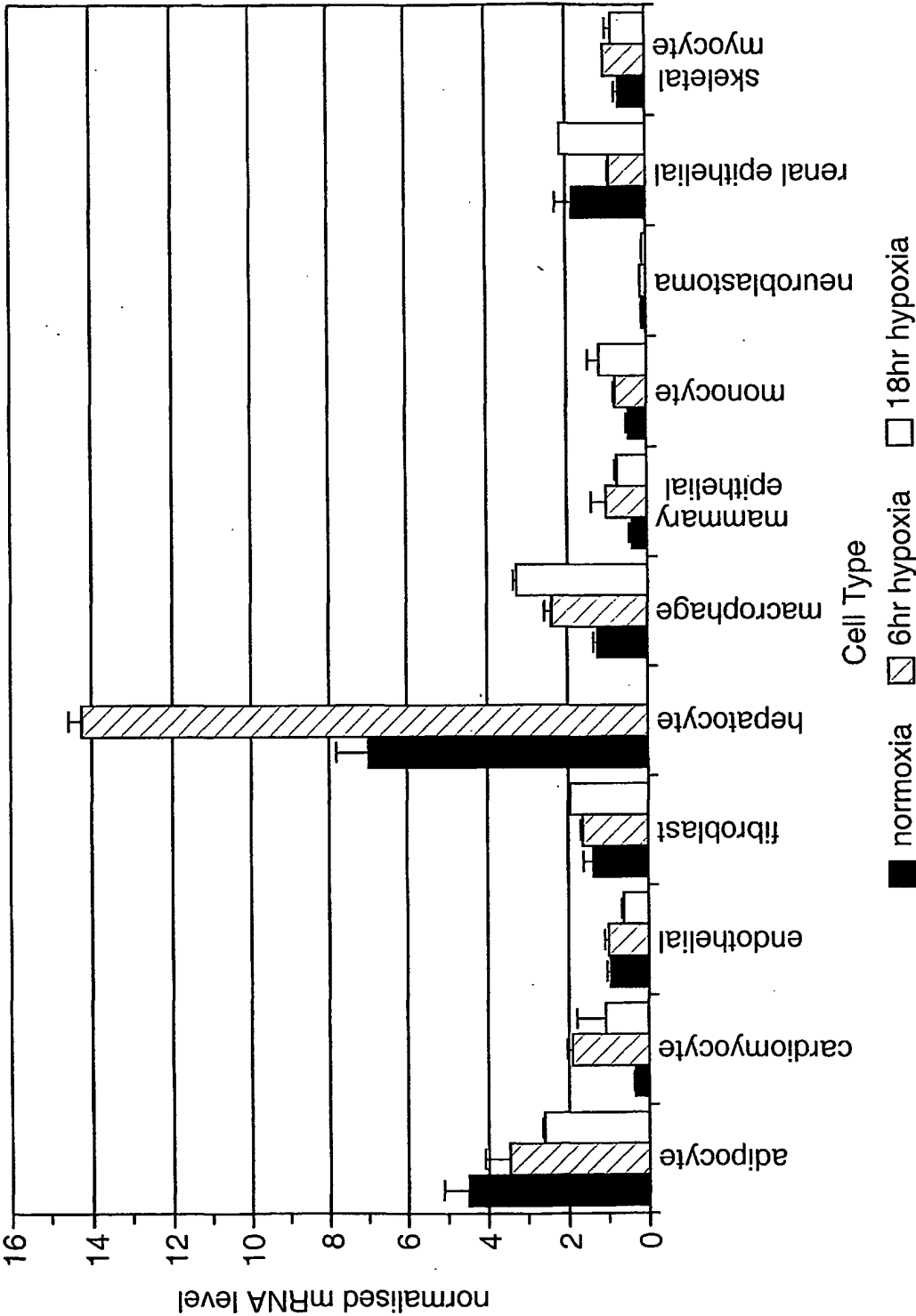
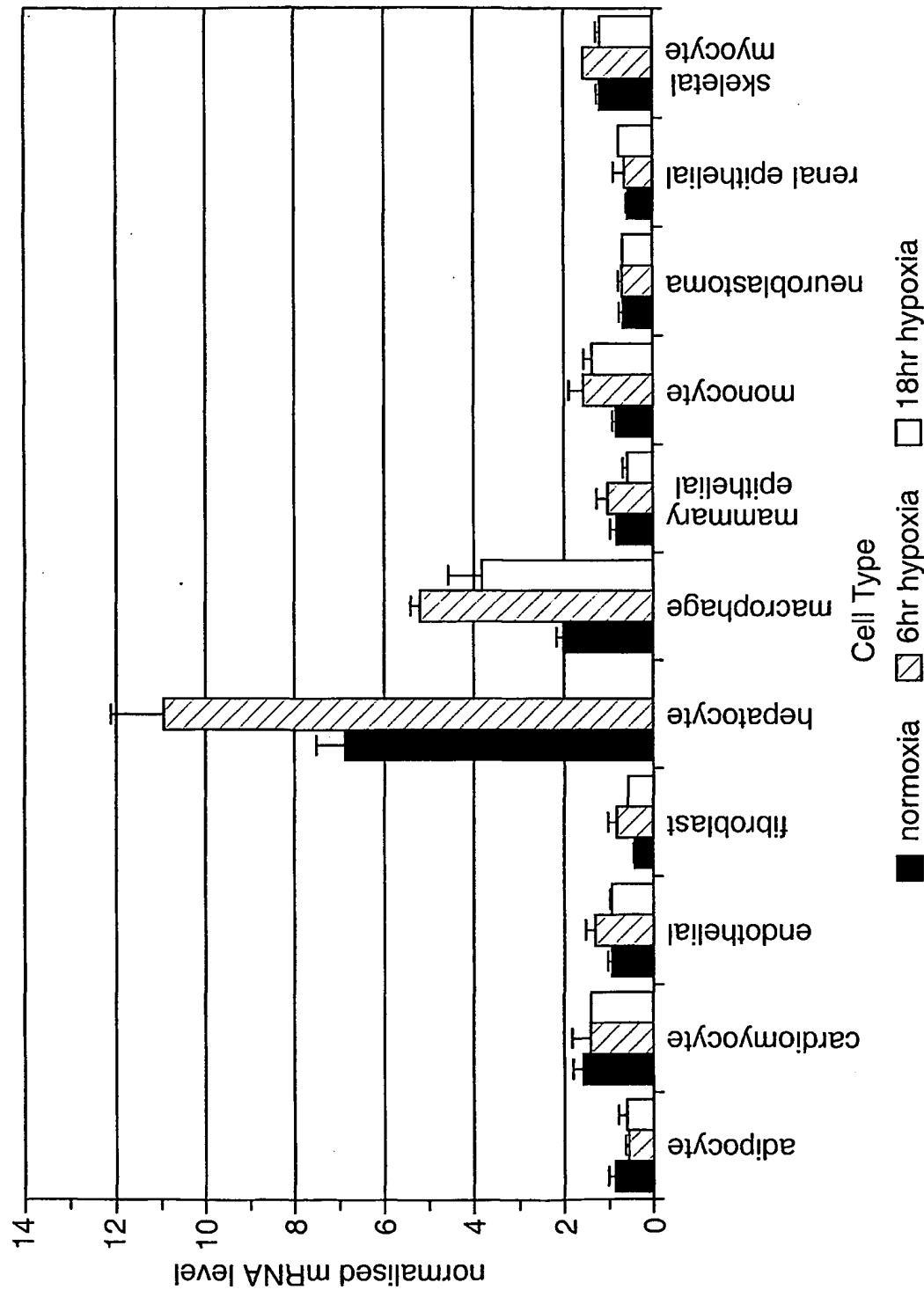
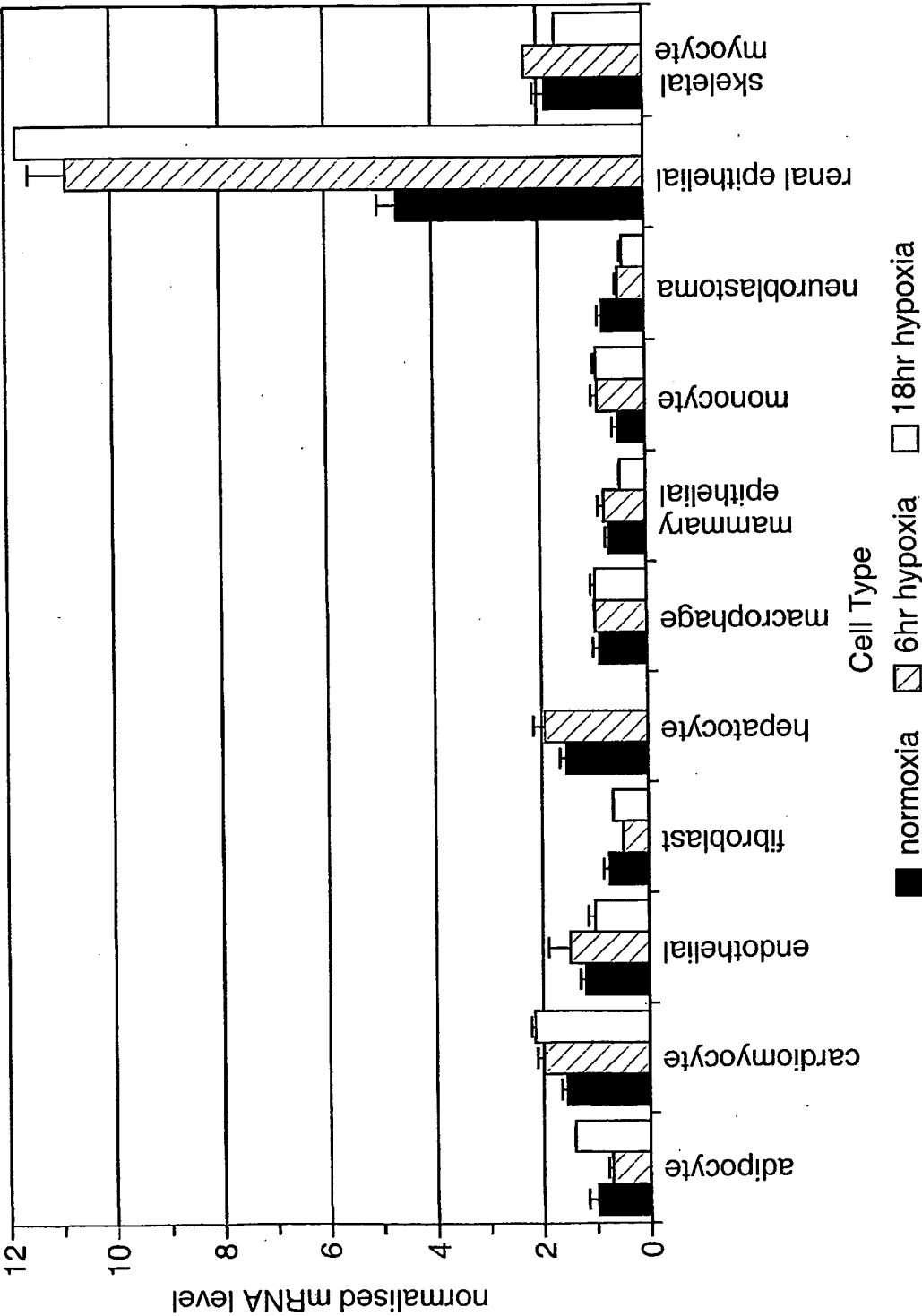


FIG. 38 p1E5/ SeqID:142/ Hepcidin antimicrobial peptide



51/96

FIG. 39 p1D24/ SeqID:118/ EST



52/96

FIG. 40 p1D21/ SeqID:130/ Hypothetical protein FLJ22622

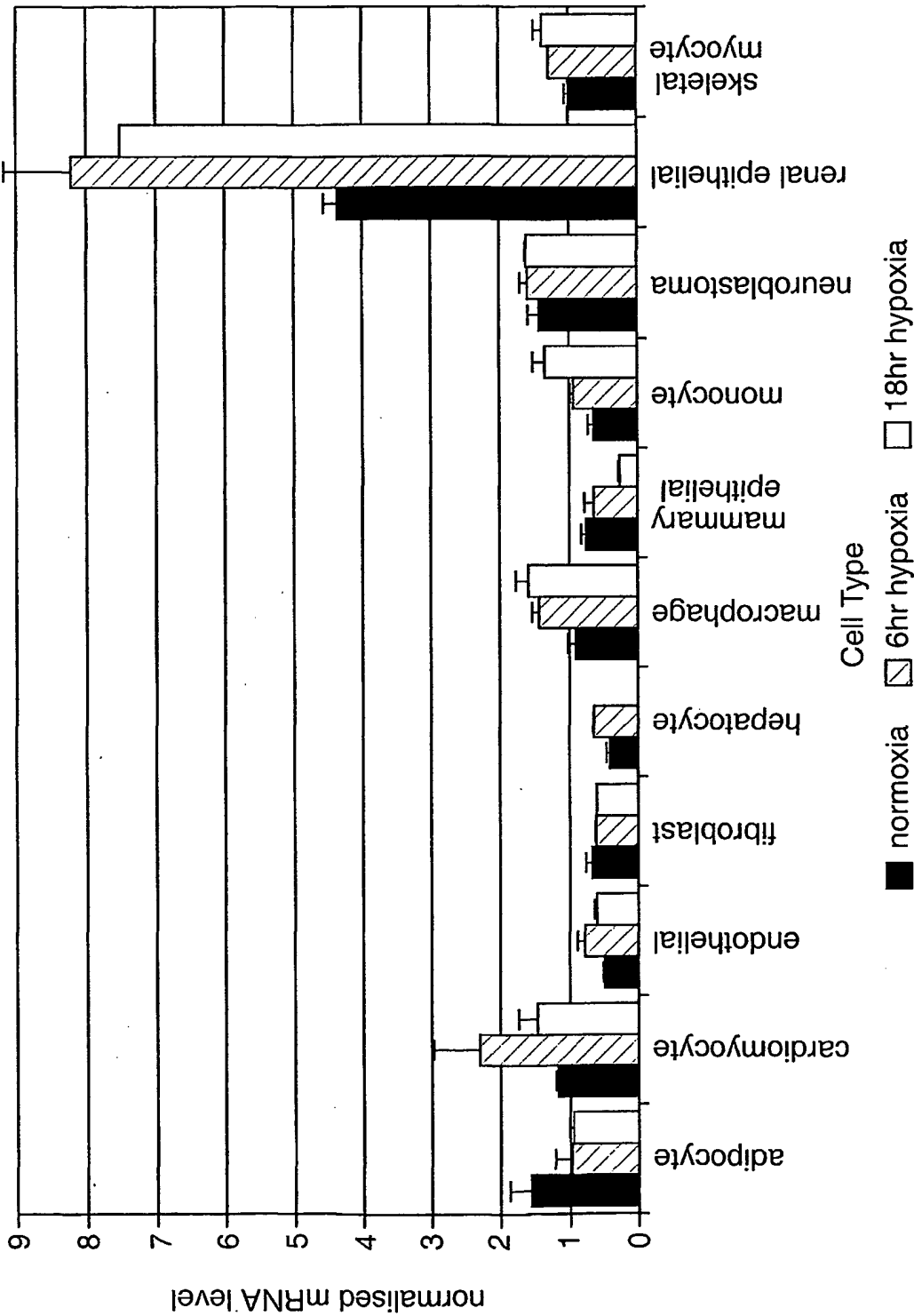


FIG. 41 p1D15/ SeqID:32/ TRIP-Br2

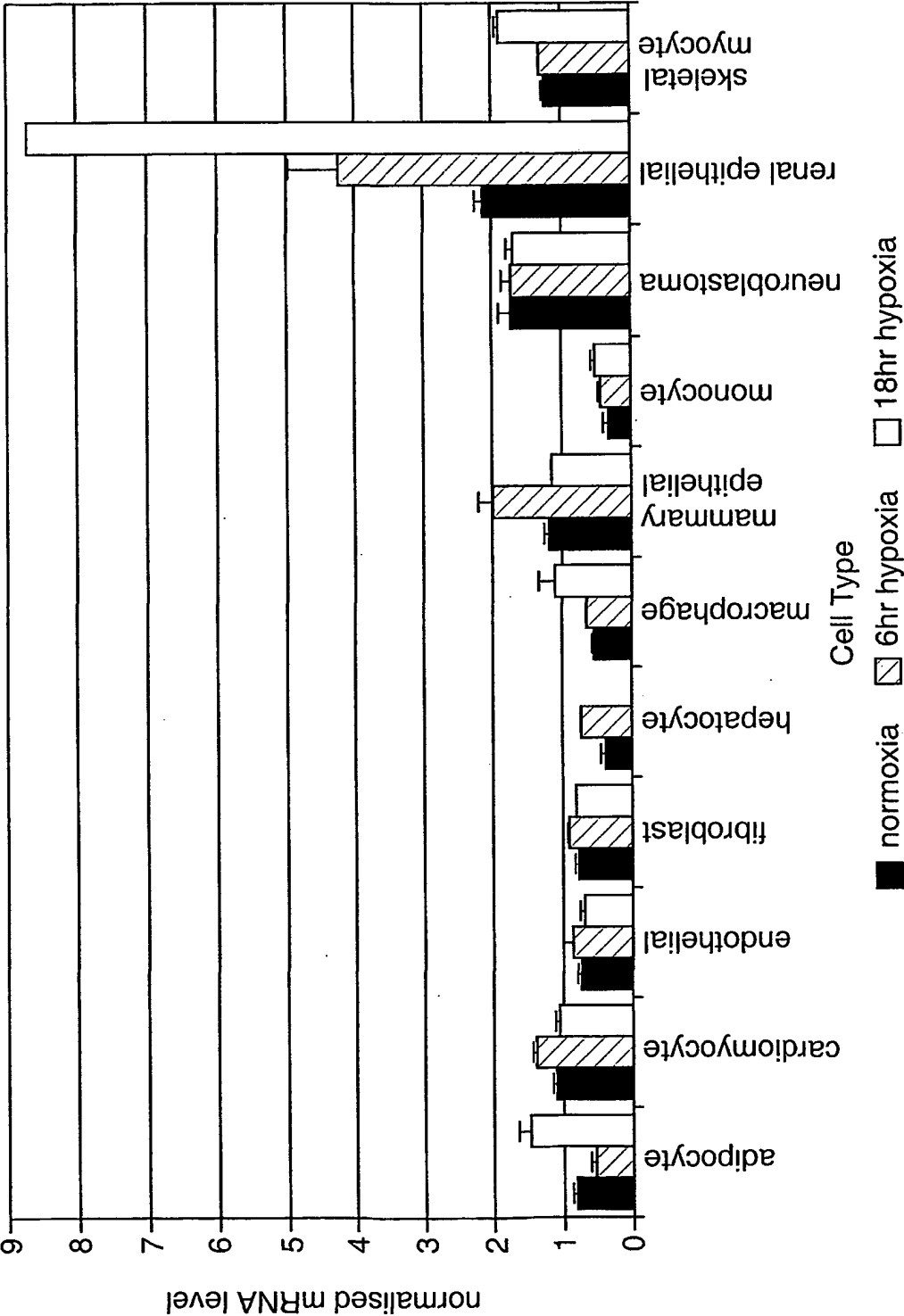
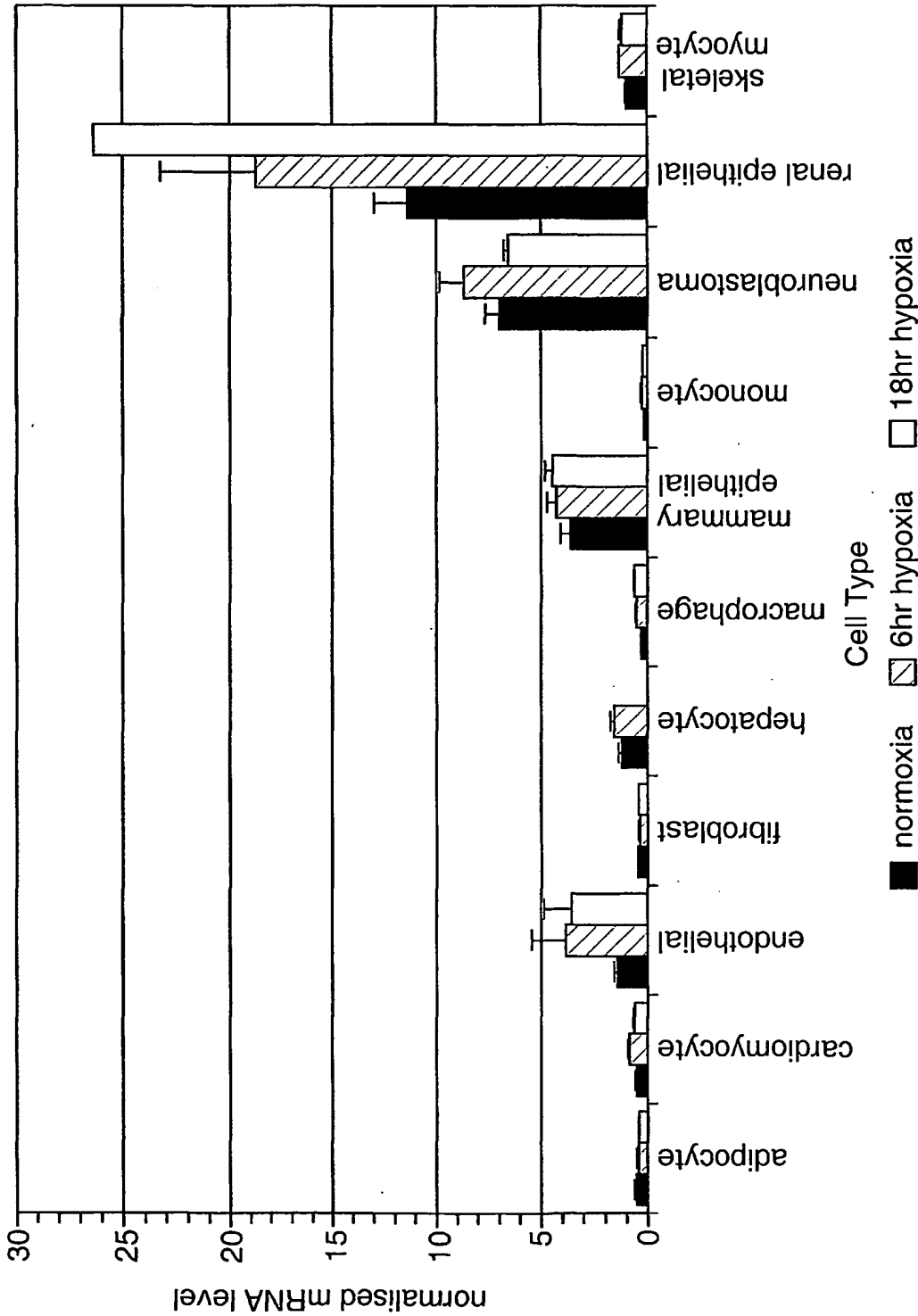
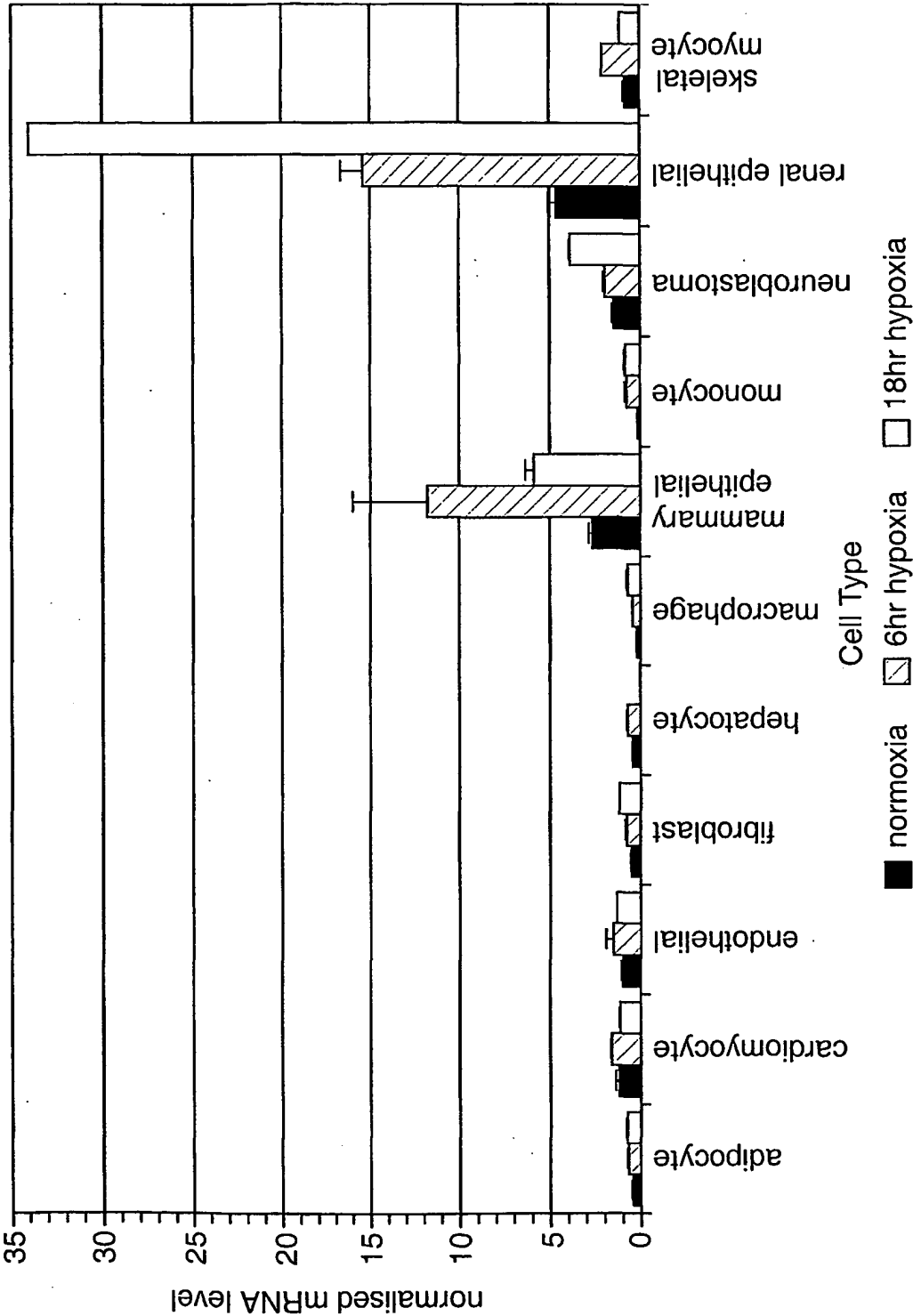


FIG. 42 p1G11/ SeqID:302/ Tumor protein D52



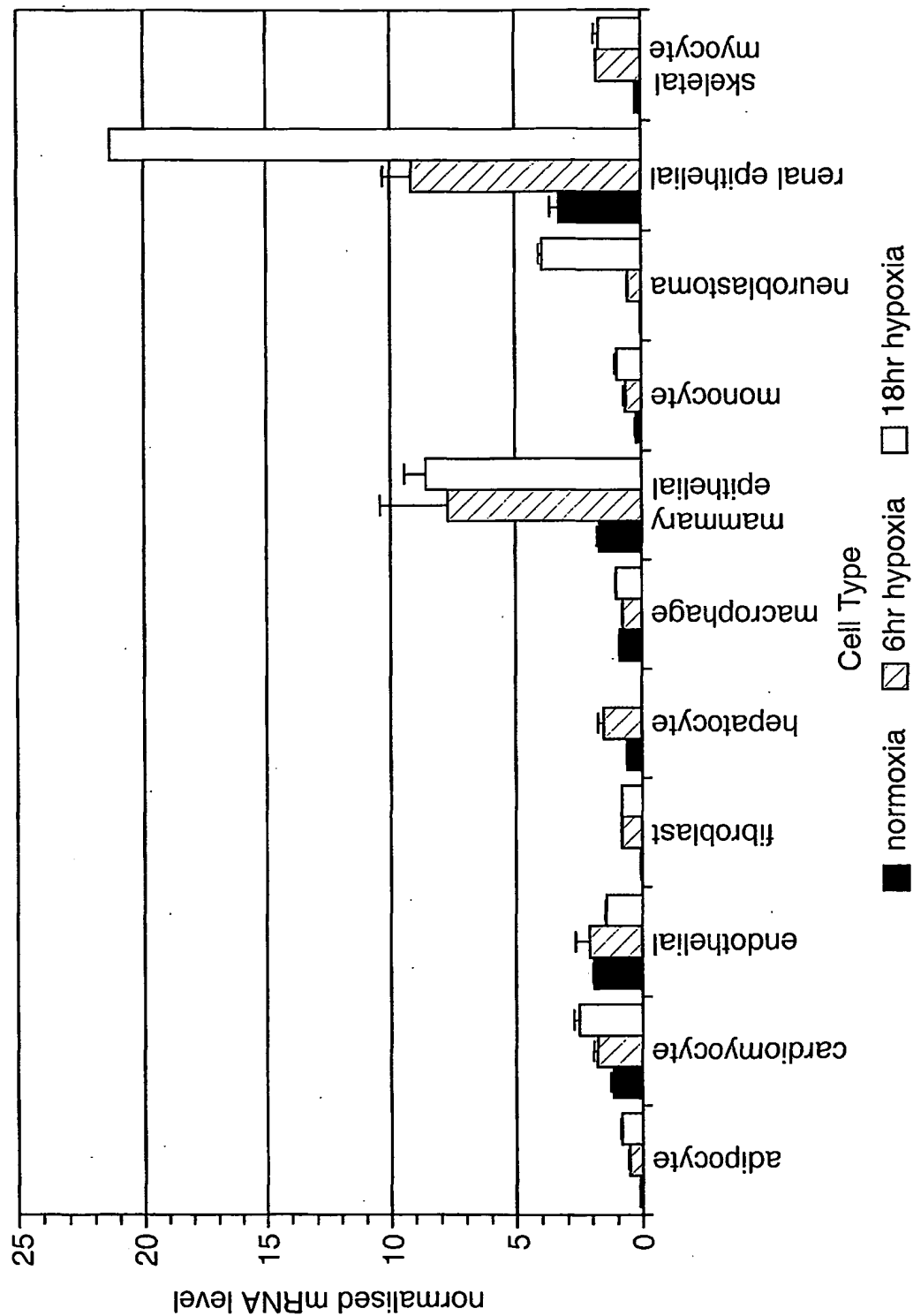
55/96

FIG. 43 p1P14/ SeqID:92/ Semaphorin 4b



56/96

FIG. 44 p1C8/ SeqID:372/ Dec-1



57/96

FIG. 45 p1J23/ SeqID:448/ Calgranulin A

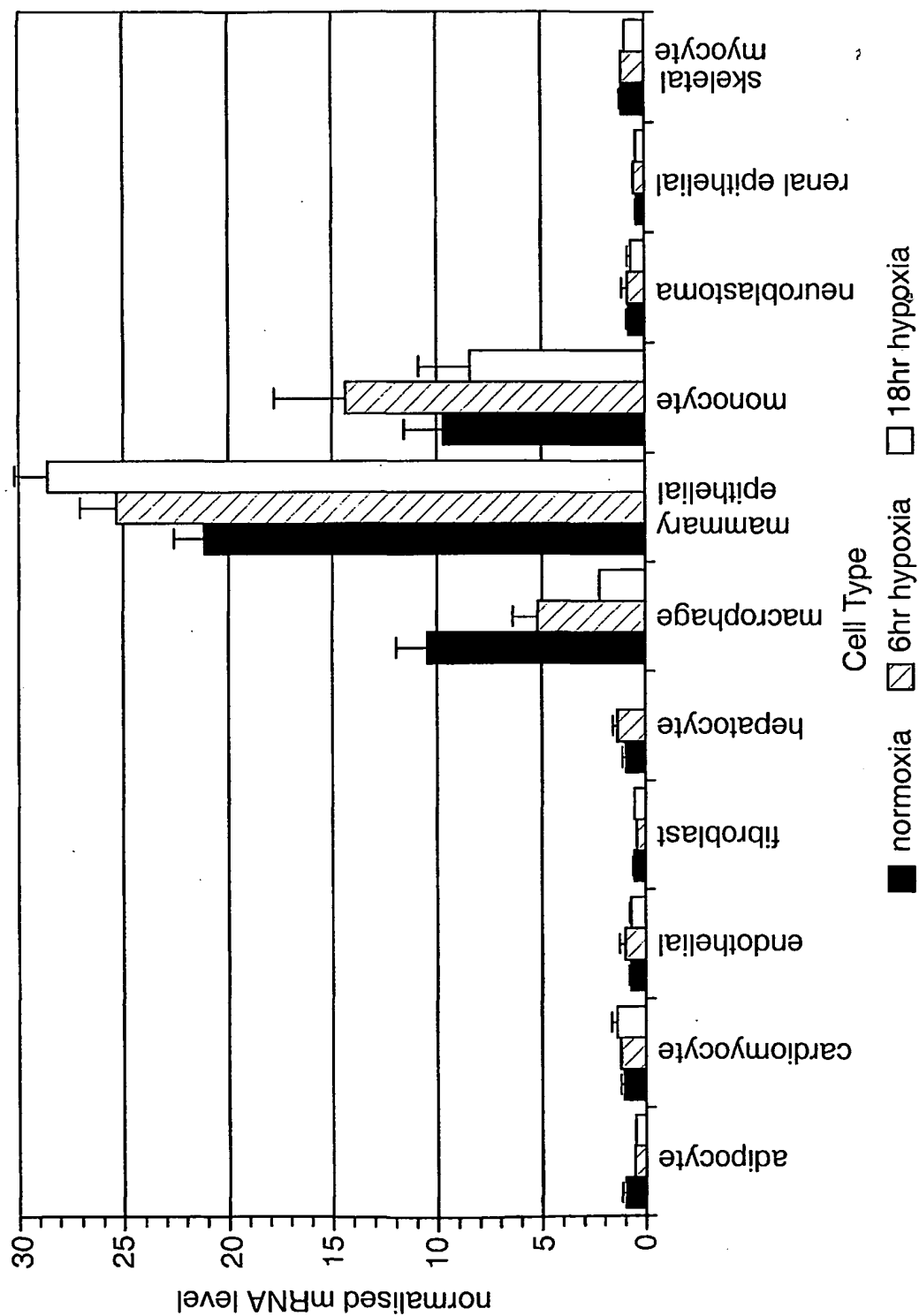
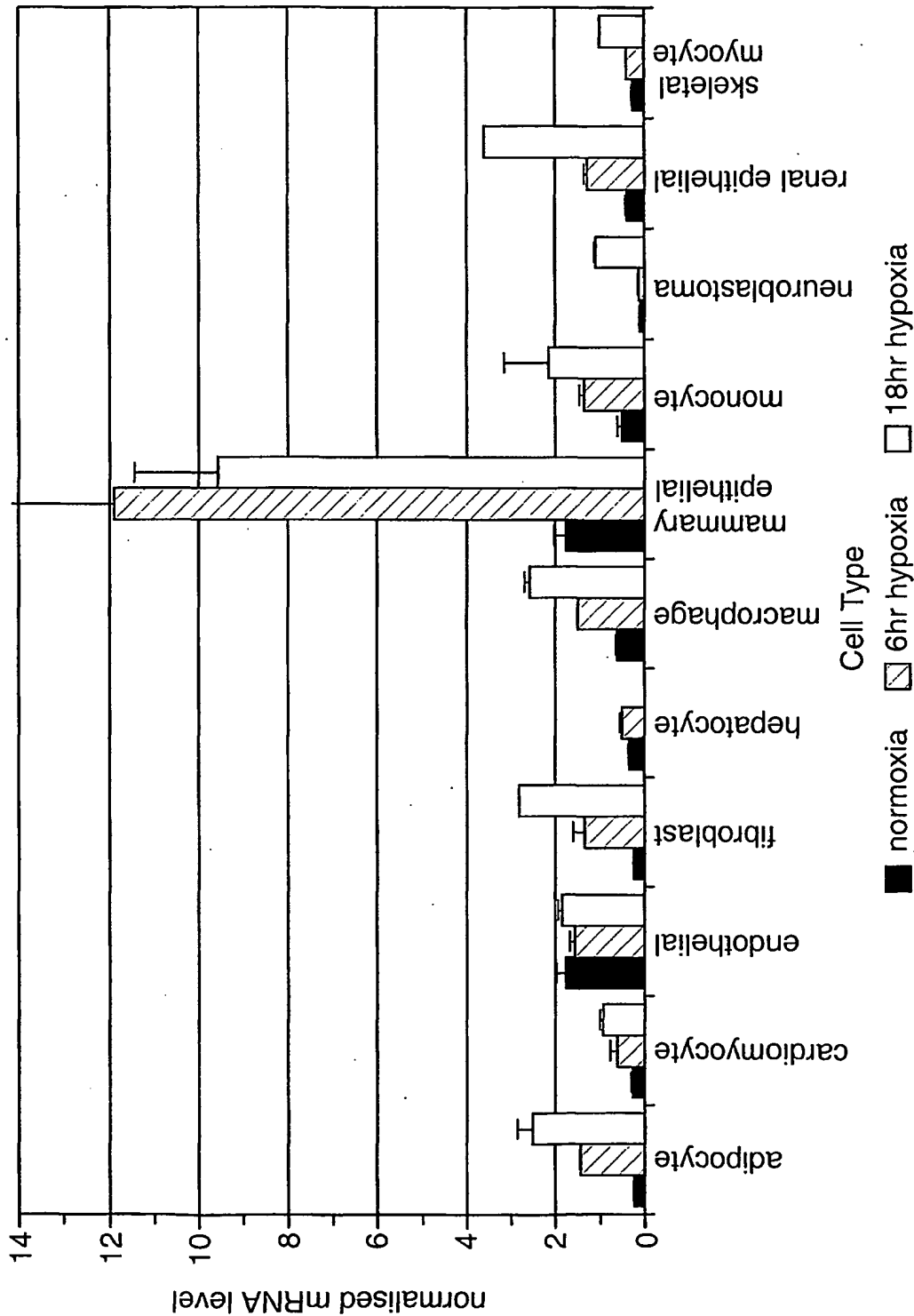
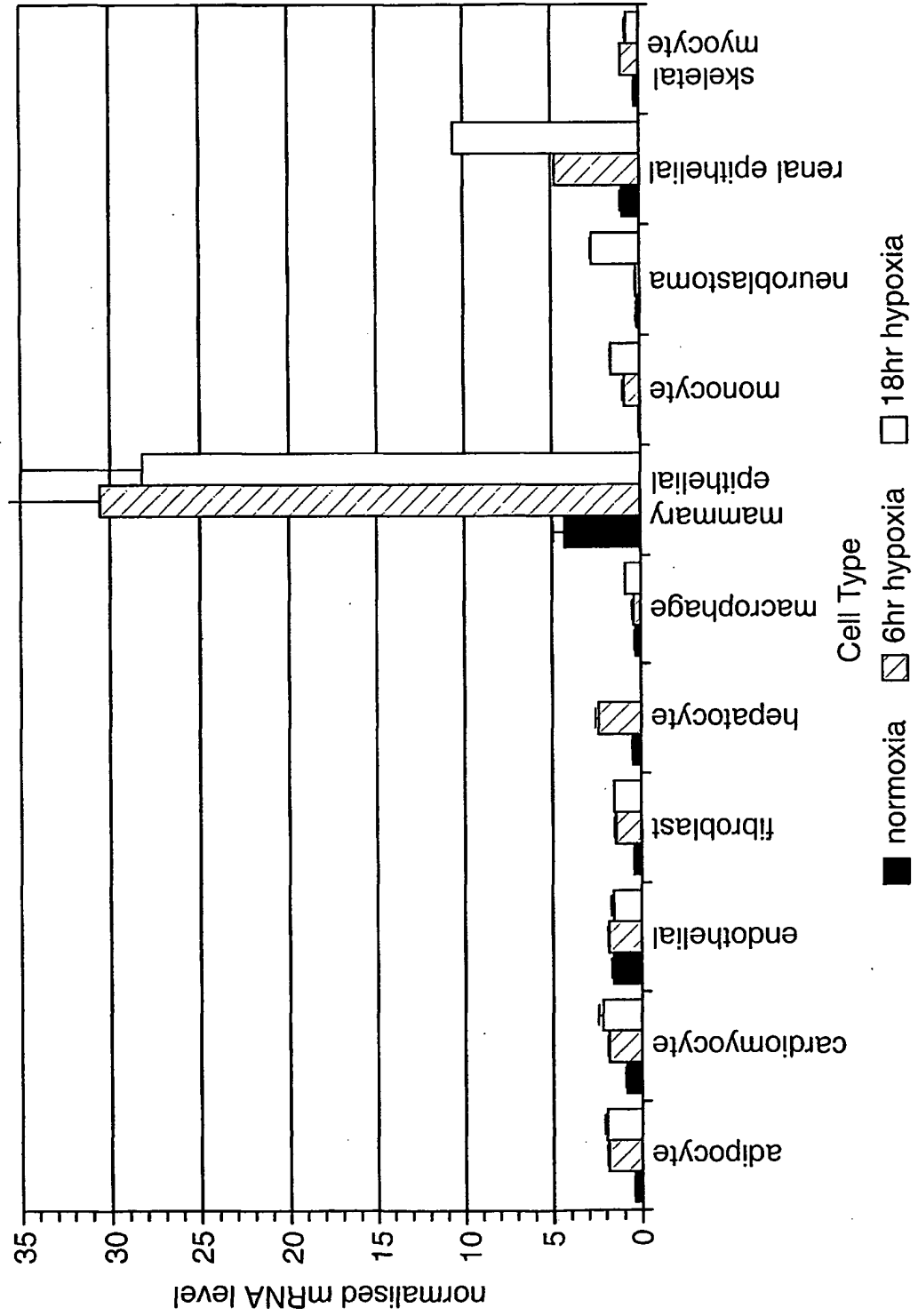


FIG. 46 p1D6/ SeqID:68/ ERO1 (*S. cerevisiae*)-like



59/96

FIG. 47 p1D4/ SeqID:26/ Hypothetical protein FLJ20500



60/96

FIG. 48 p1B2/ SeqID:230/ N-myc downstream regulated

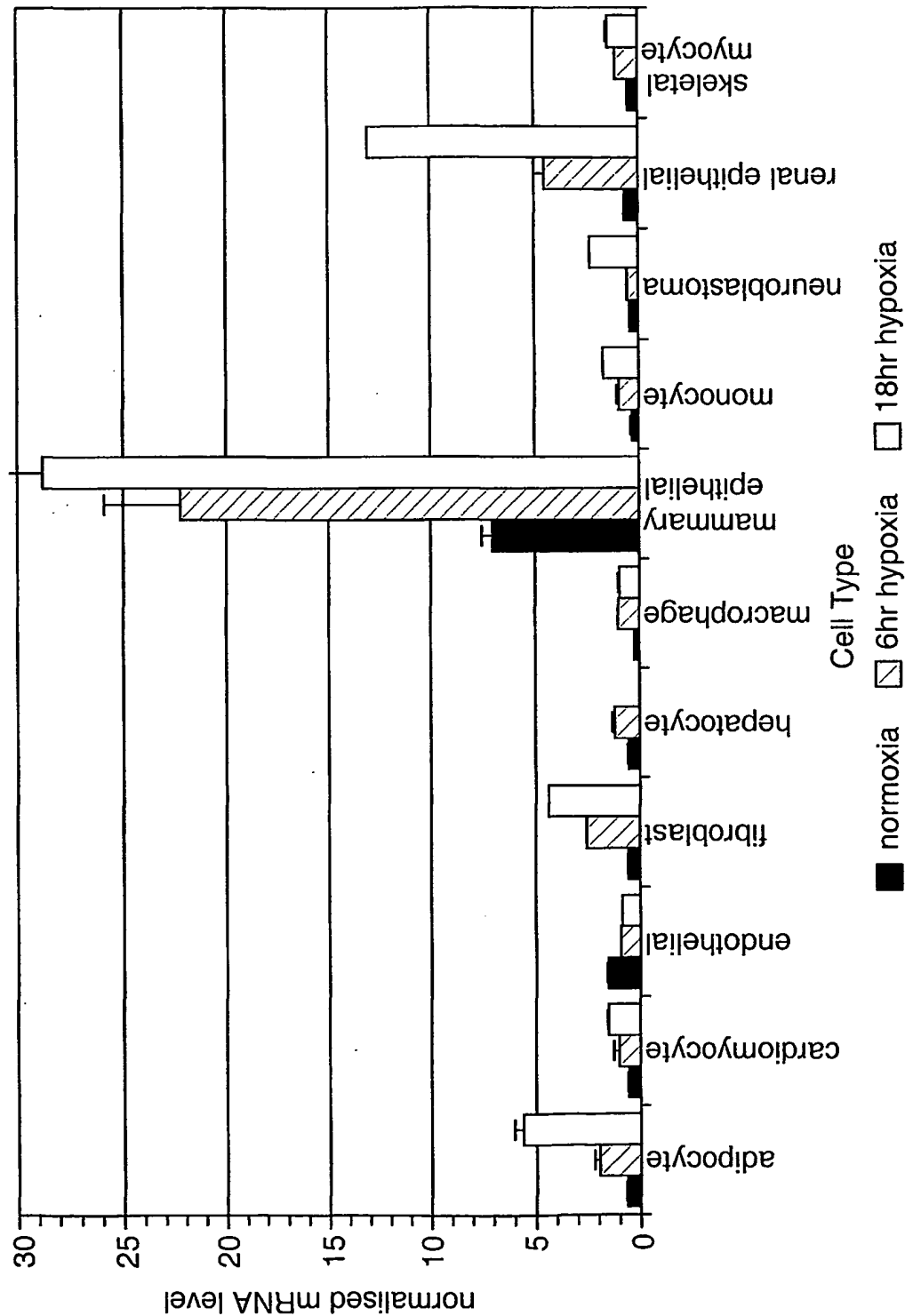
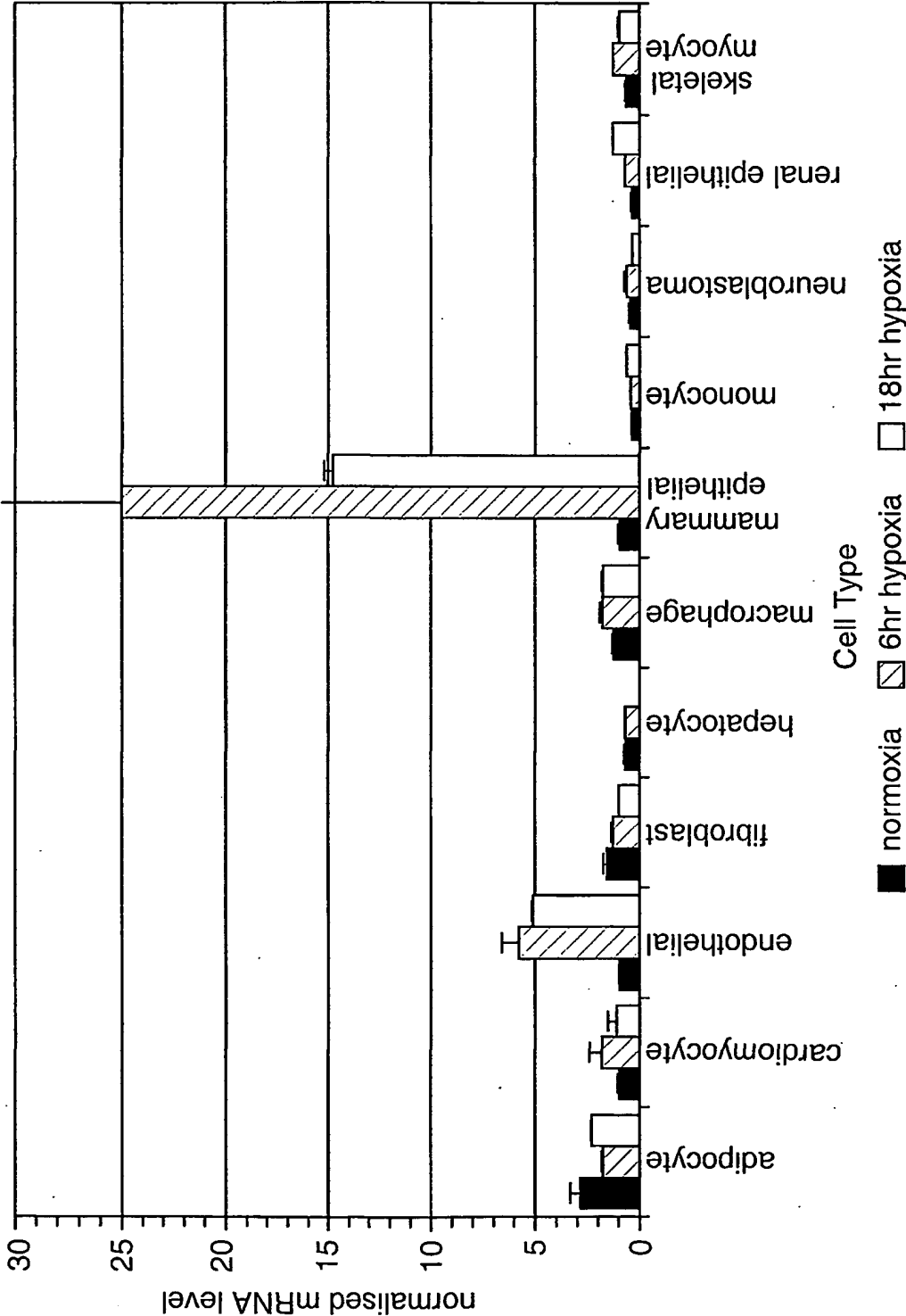
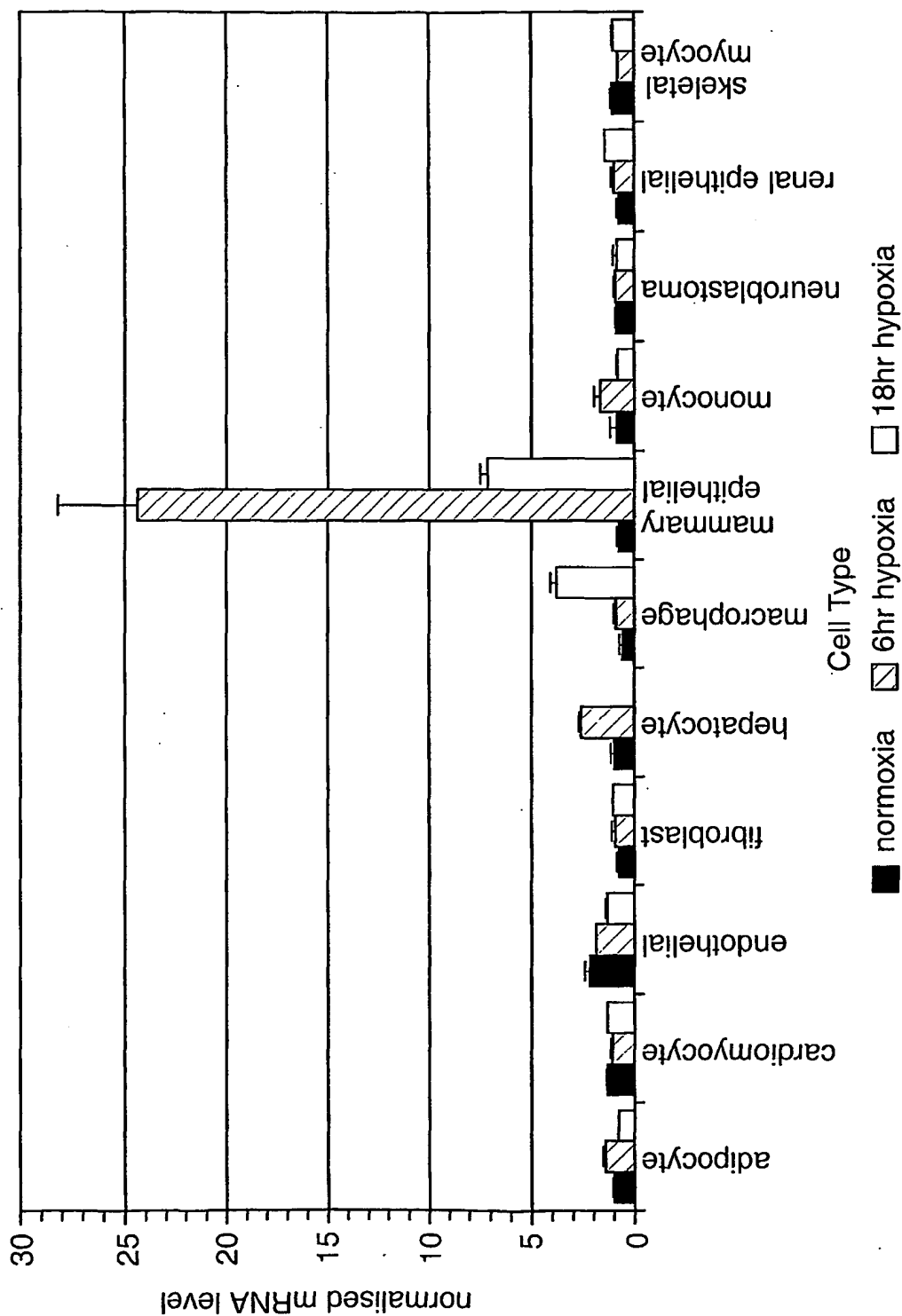


FIG. 49_{p1C16/ SeqID:388/} Decidual protein induced by progesterone



62/96

FIG. 50 p1C12/ SeqID:380/ Integrin, alpha 5



63/96

FIG. 51 p1B17/ SeqID:226/ Tissue factor

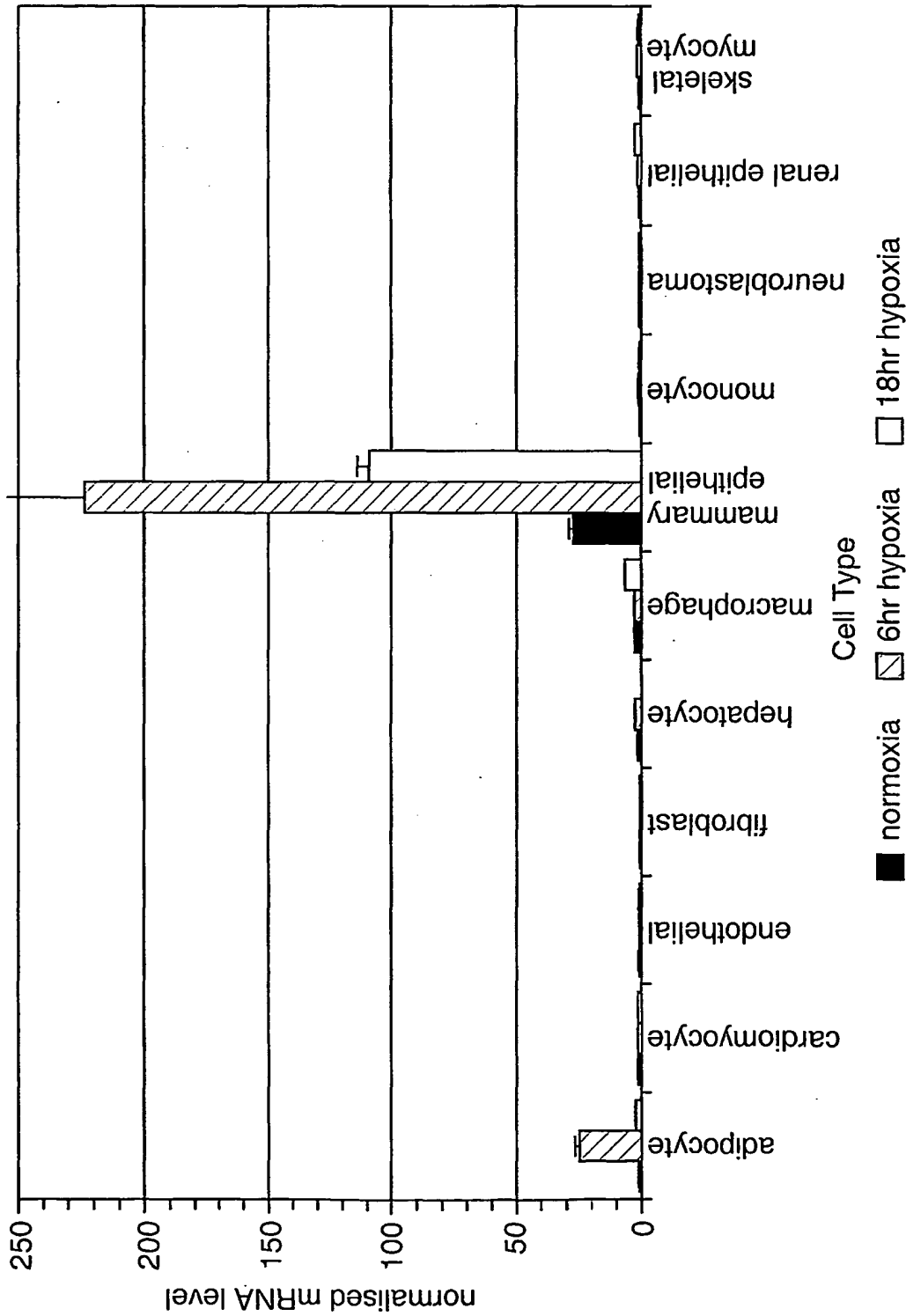
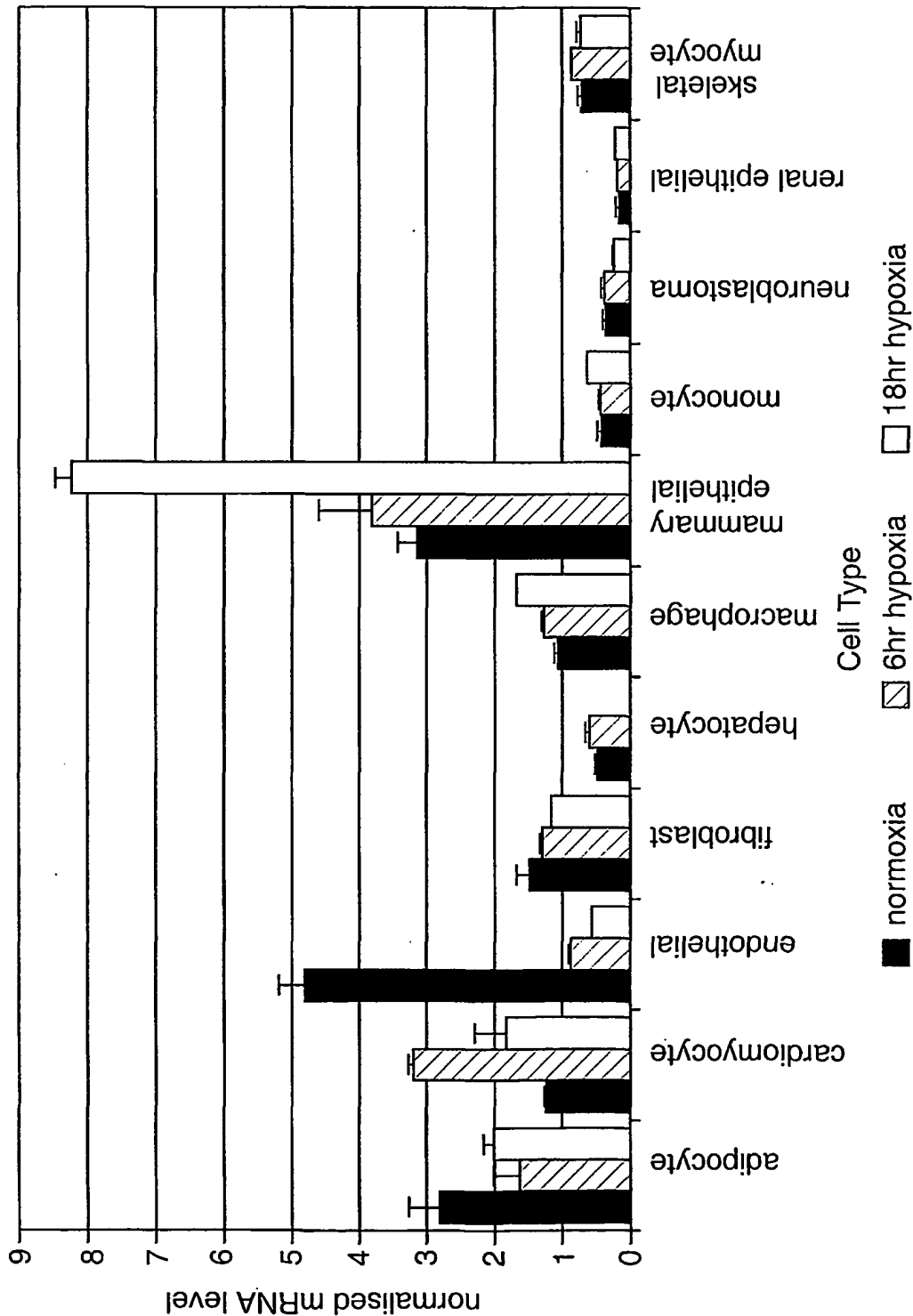


FIG. 52 p1N17/ SeqID:238/ COX-2

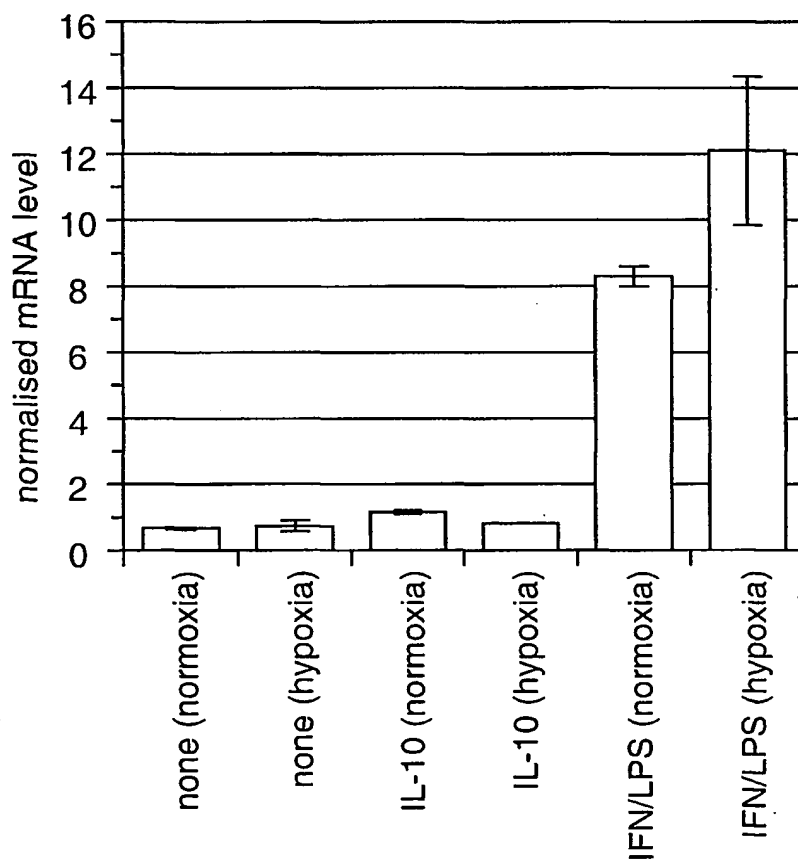


65/96

FIG. 53a

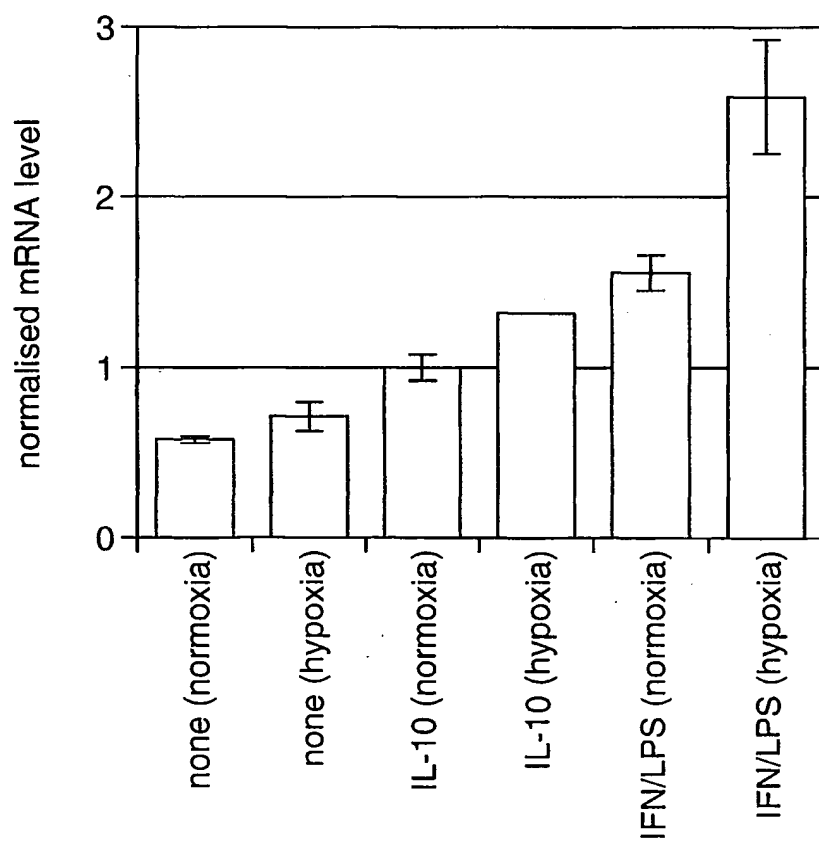
p1E10/ SeqID:72

cDNA FLJ11041 fis, clone PLACE1004405



66/96

FIG. 53b

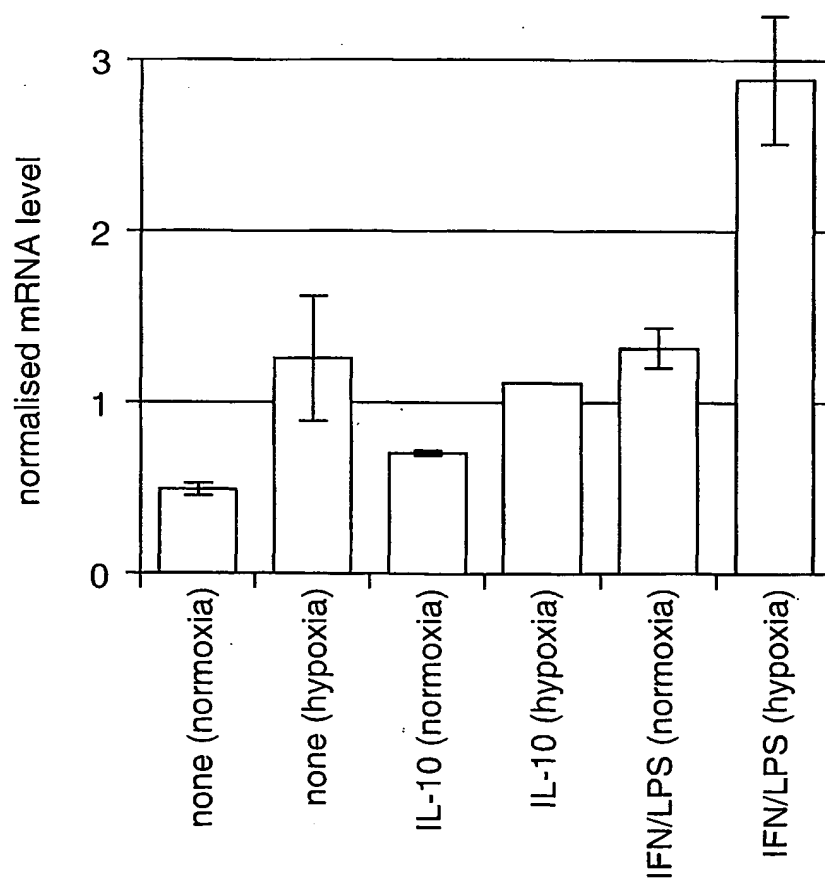
p1D24/ SeqID:118
EST

67/96

FIG. 53c

p1E7/ SeqID:84

Novel metallothionein

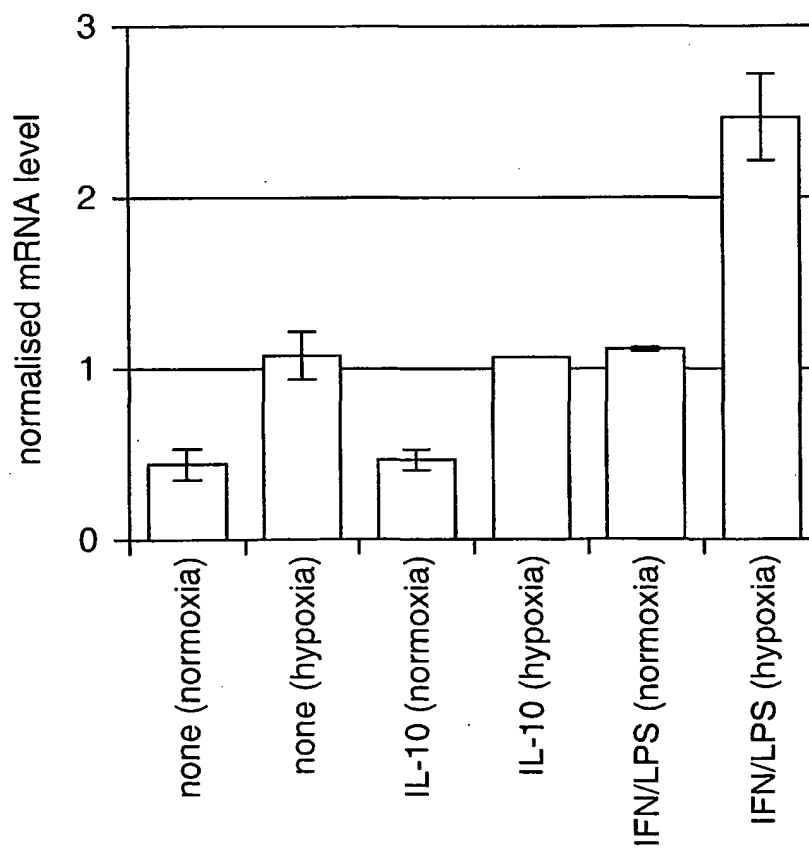


68/96

FIG. 53d

p1F6/ SeqID:338

Hypothetical protein hqp0376

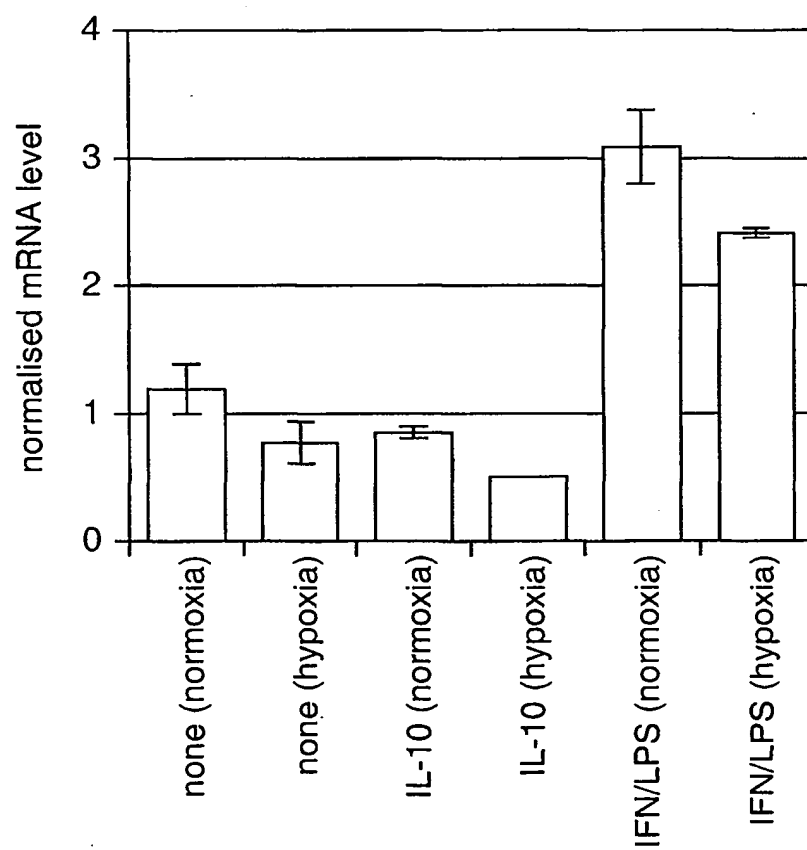


69/96

FIG. 53e

p1E22/ SeqID:162

cDNA FLJ13618 fis, clone PLACE1010925

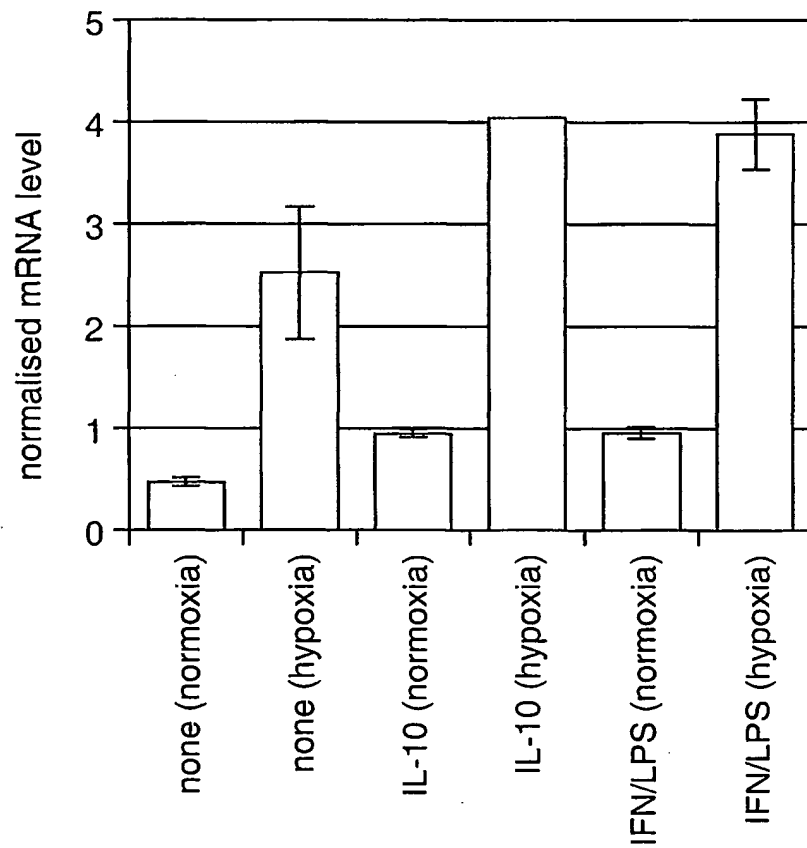


70/96

FIG. 53f

p1P14/ SeqID:92

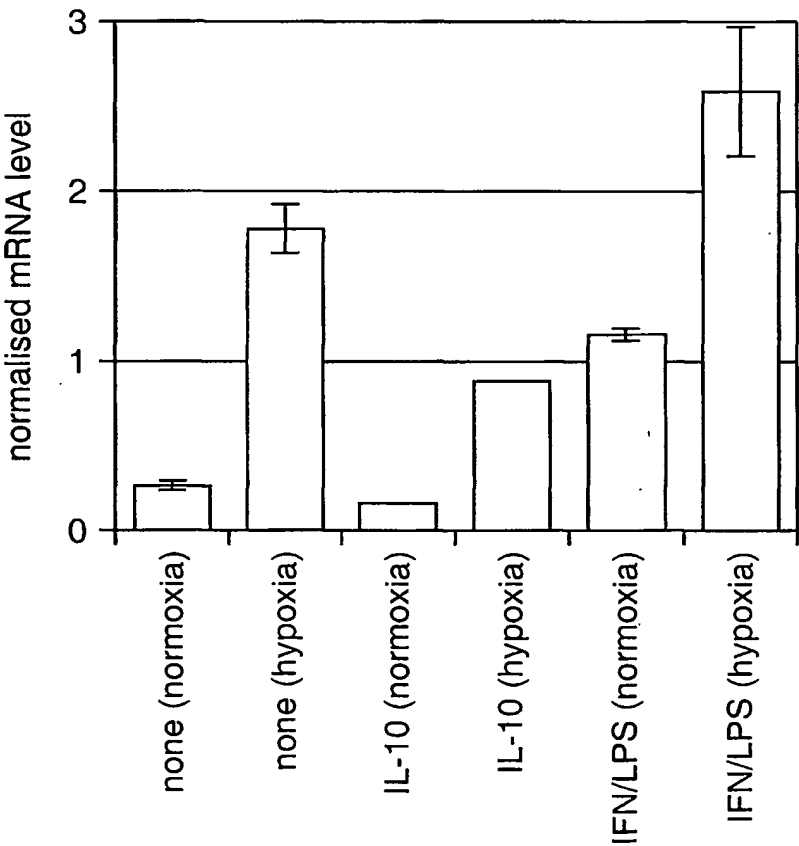
Semaphorin 4b



71/96

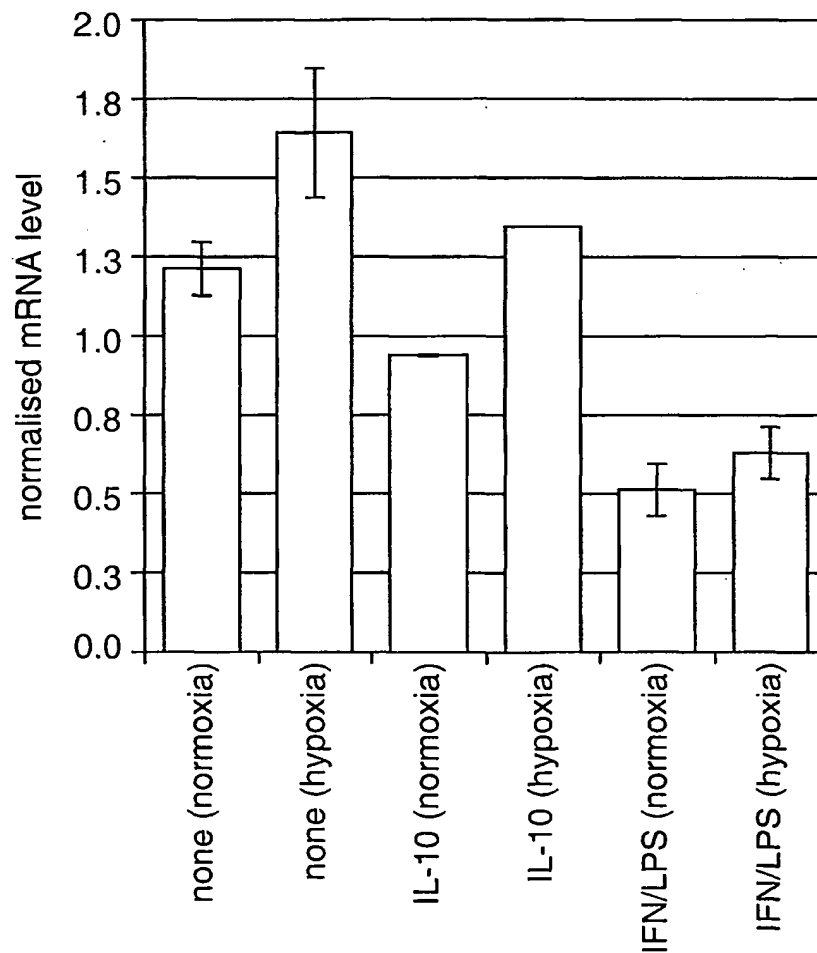
FIG. 53g

p1F17/ SeqID:330
P8 protein (candidate of metastasis 1)



72/96

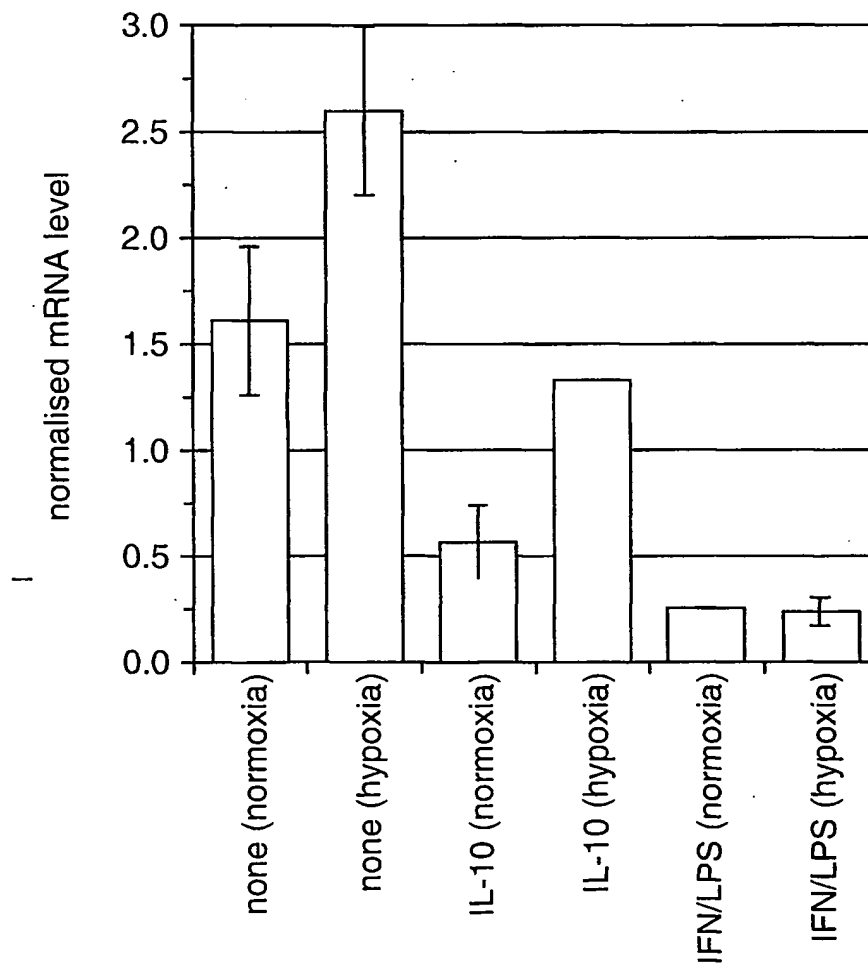
FIG. 54a

p1E1/ SeqID:124
EST

73/96

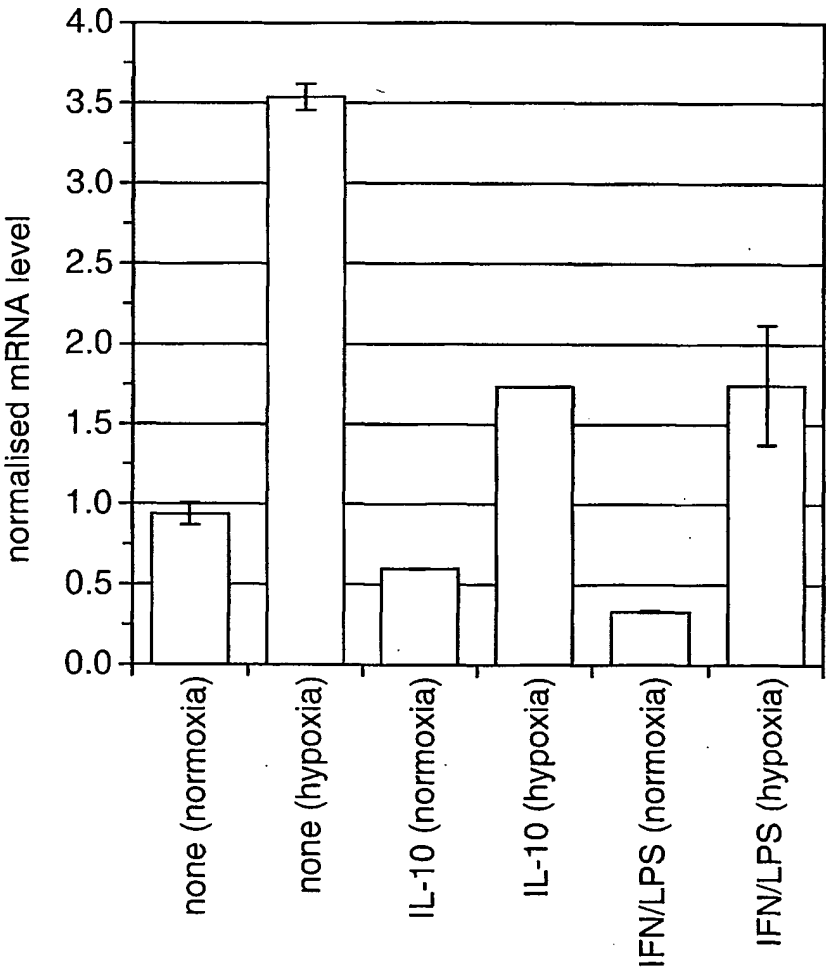
FIG. 54b

p1D18/ SeqID:128
cDNA FLJ13443 fis, clone PLACE1002853



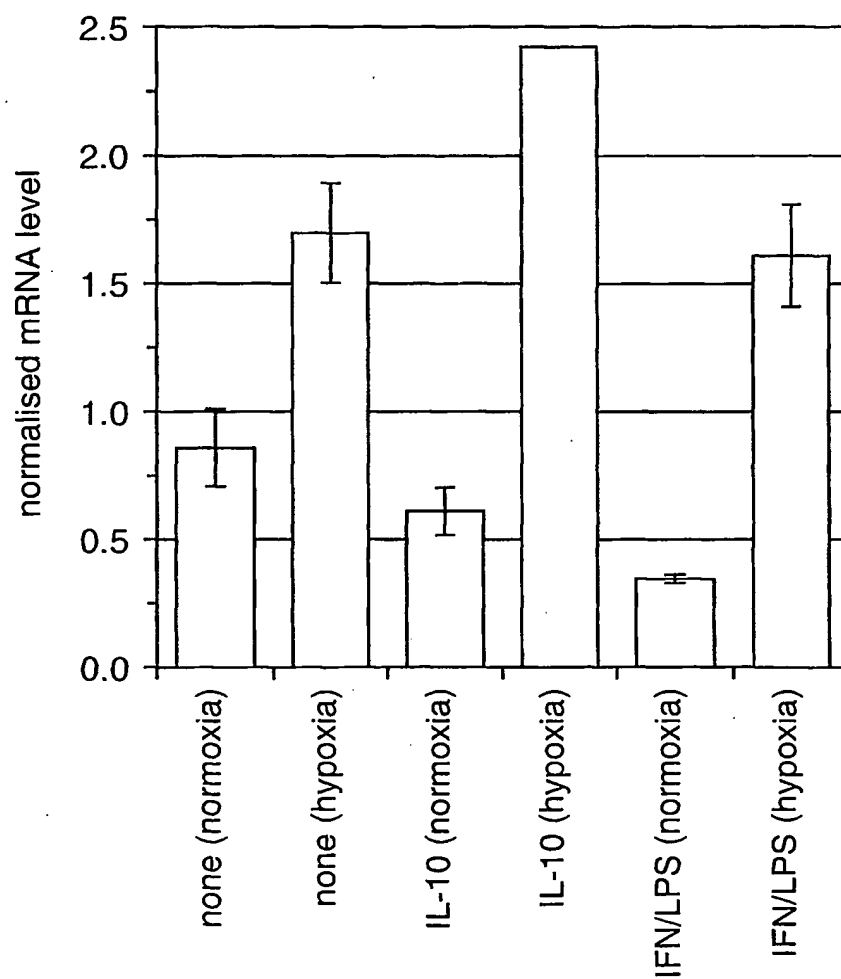
74/96

FIG. 54c
p1F9/ SeqID:20
Hypothetical protein KIAA0742



75/96

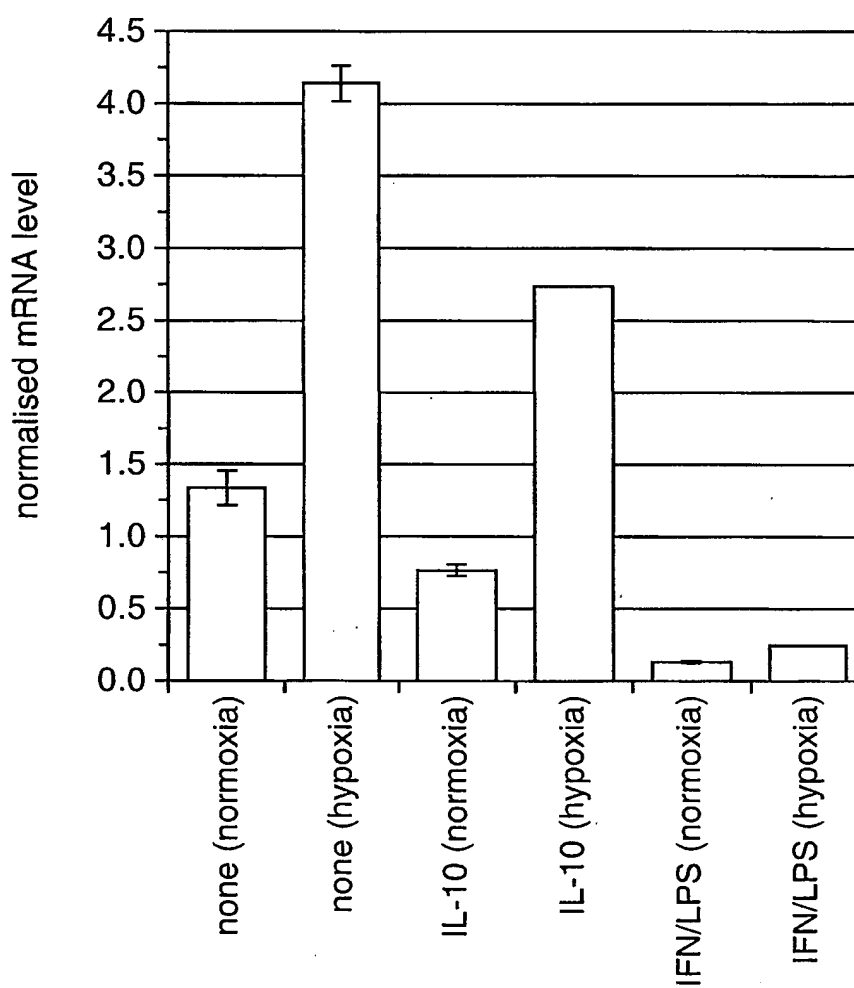
FIG. 54d
p1D1/ SeqID:24
Hypothetical protein FLJ10134



76/96

FIG. 54e

p1F8/ SeqID:10
Hypothetical protein KIAA0914

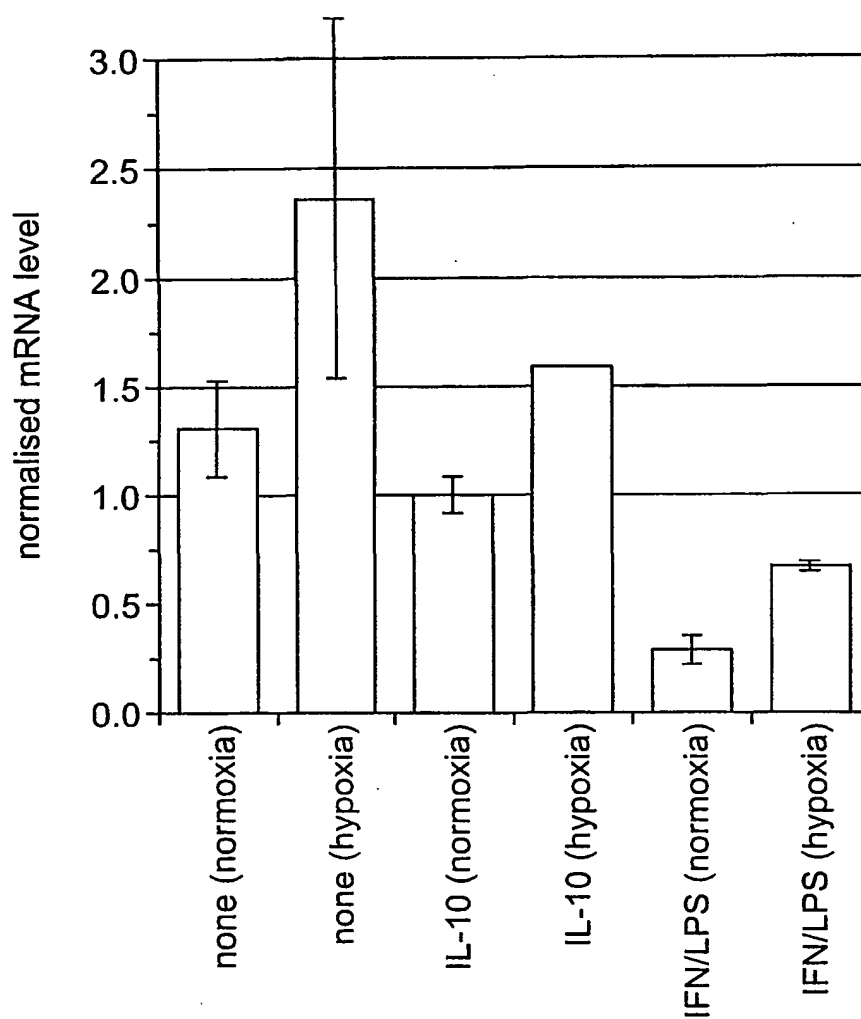


77/96

FIG. 54f

p1D16/ SeqID:34

Hypothetical protein FLJ20308

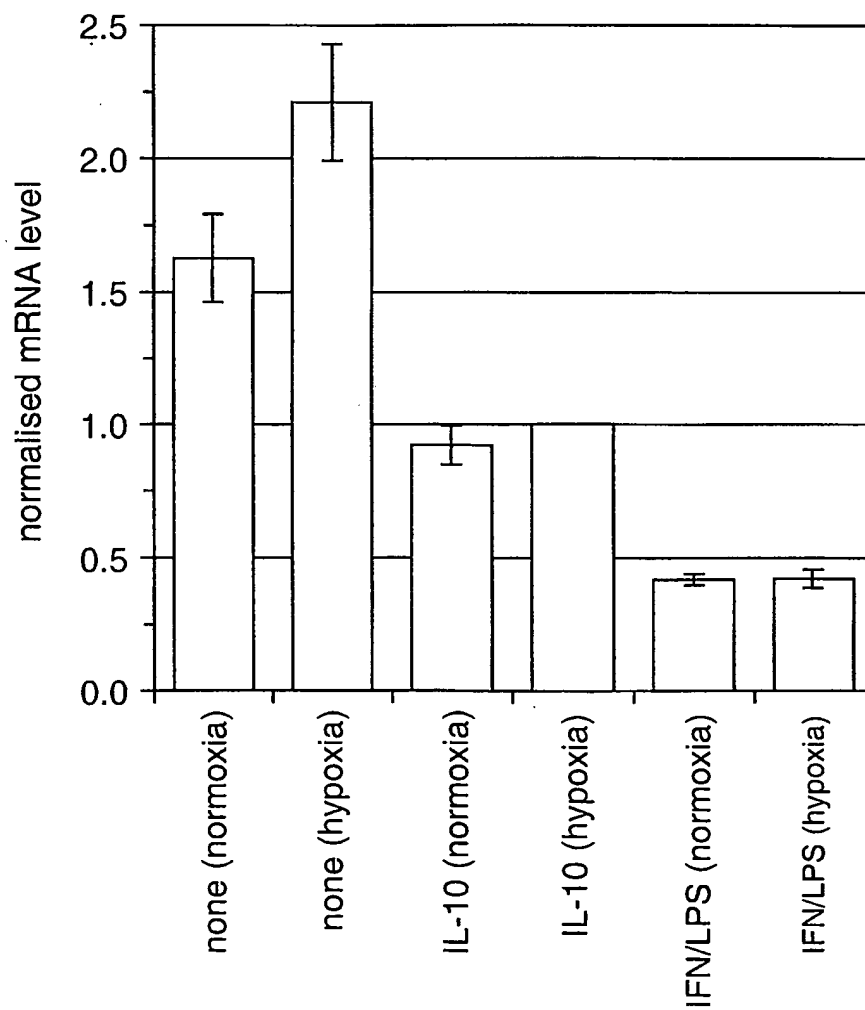


78/96

FIG. 54g

p1F3/ SeqID:334

Hypothetical protein XP_017131

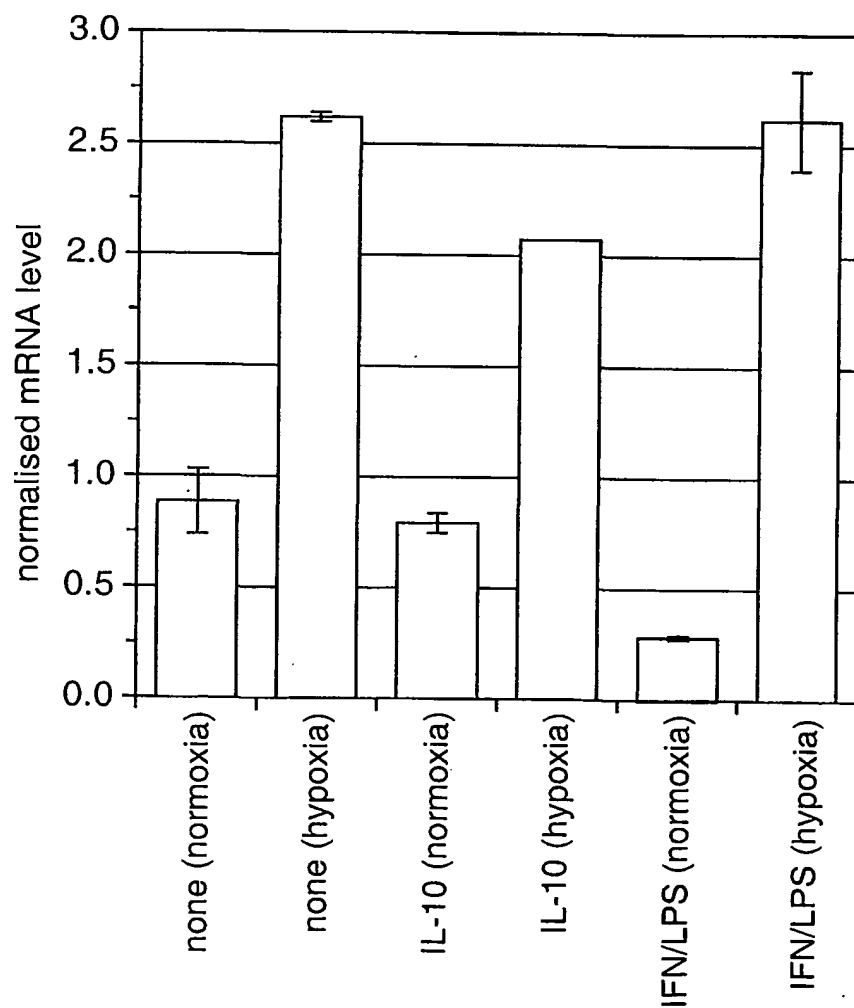


79/96

FIG. 54h

p1D12/ SeqID:30

Hypothetical protein KIAA1376

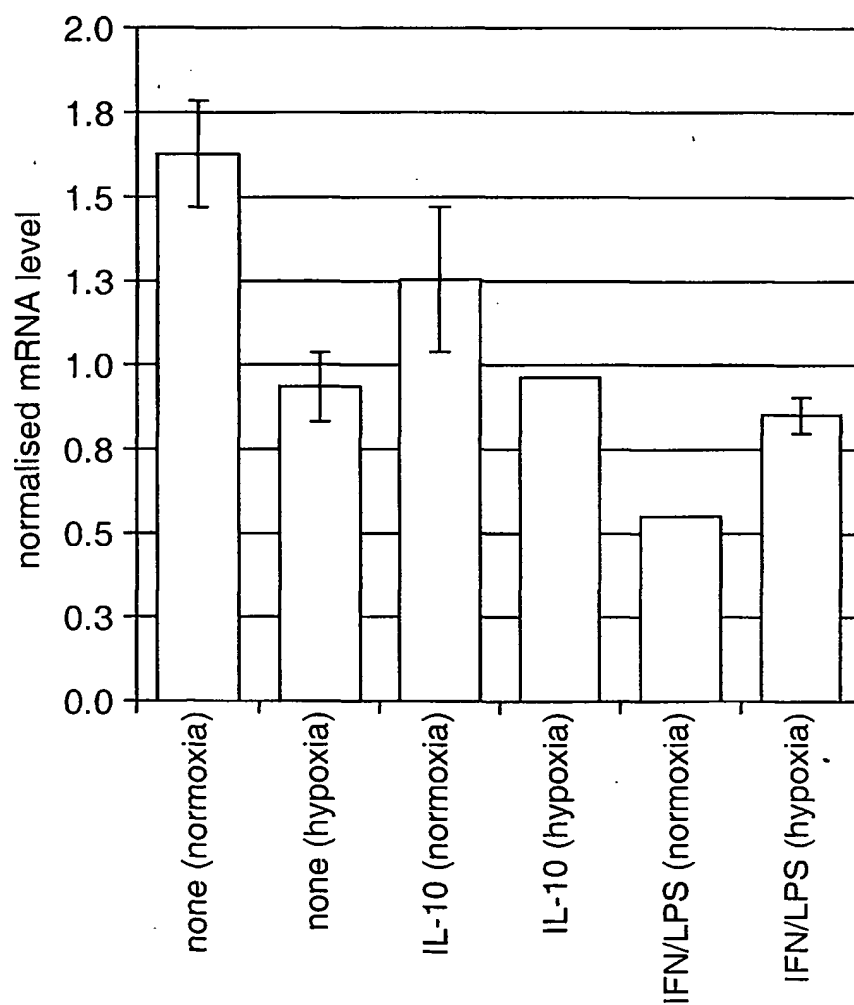


80/96

FIG. 55a

p1D9/ SeqID:28

Hypothetical protein DKFZP564D116

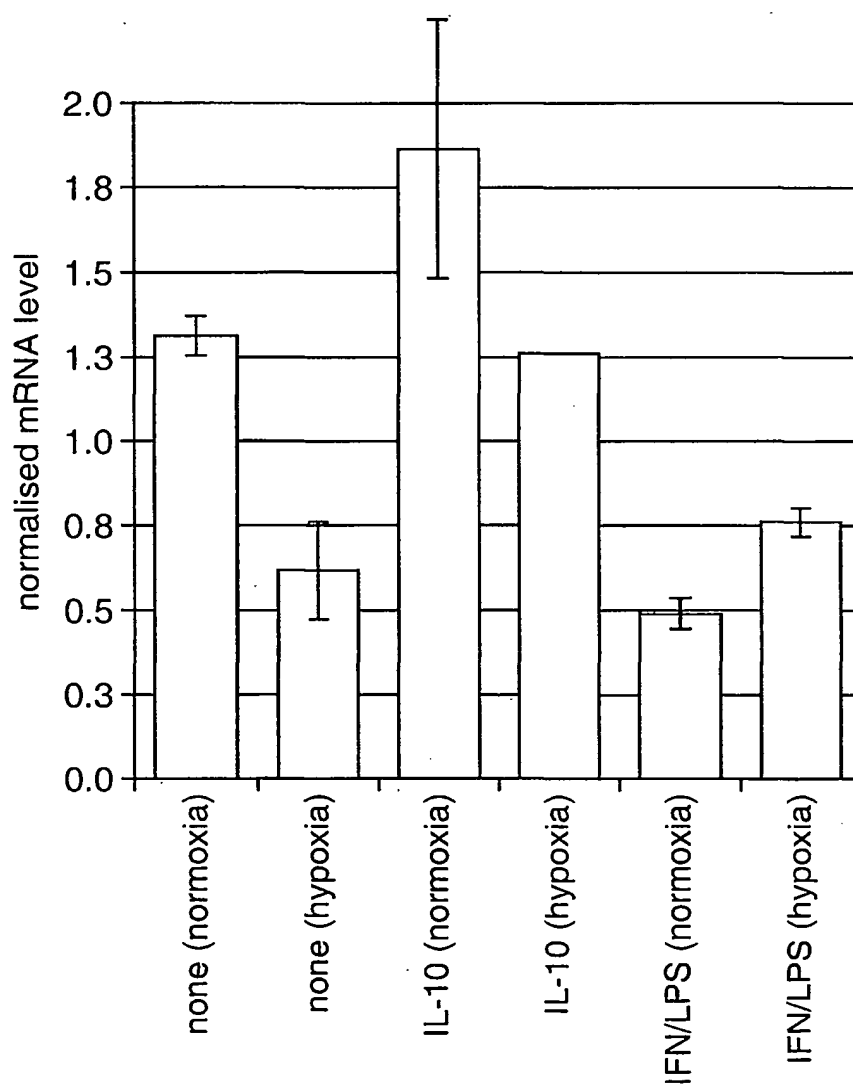


81/96

FIG. 55b

p1115/ SeqID:48

Hypothetical protein CGI-117

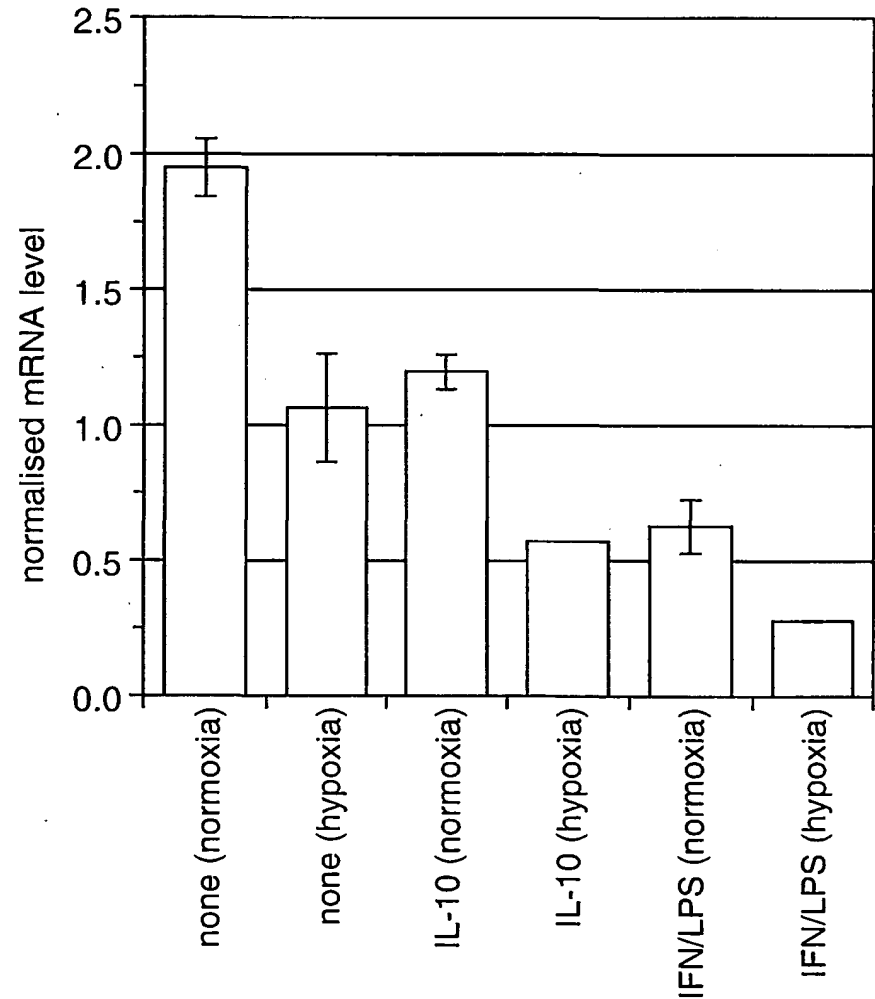


82/96

FIG. 55c

p114/ SeqID:54

Hypothetical protein HSPC196

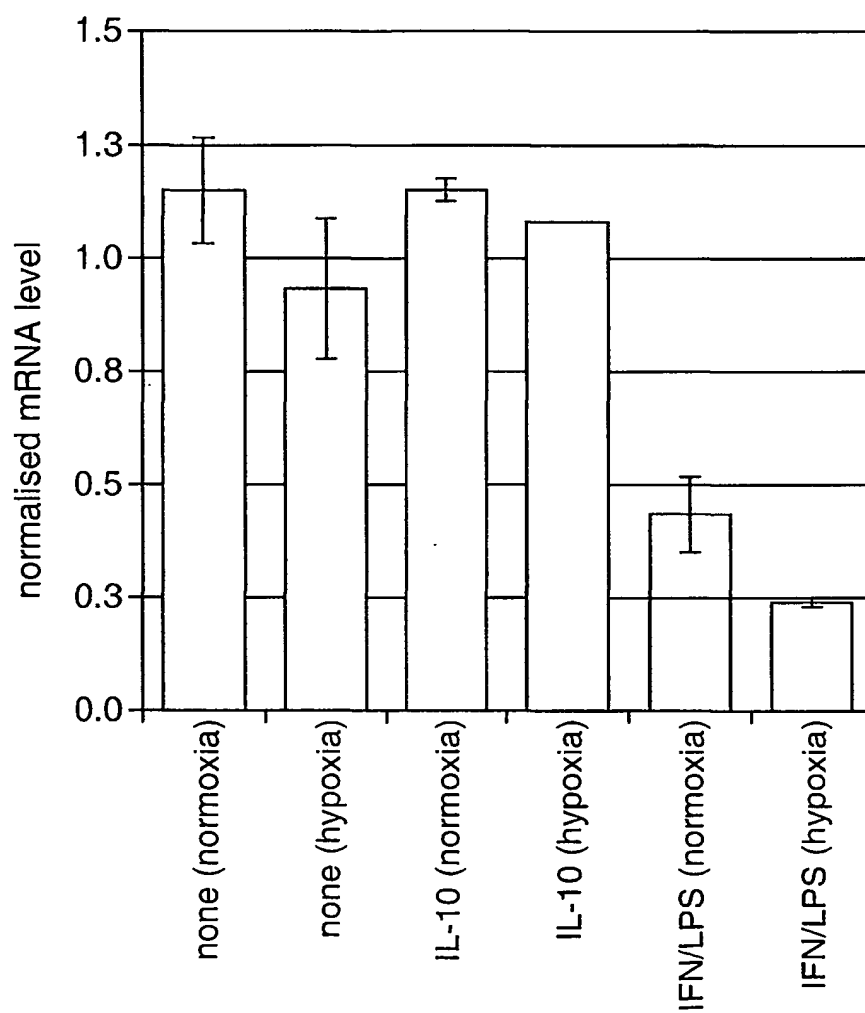


83/96

FIG. 55d

p1E13/ SeqID:22

Hypothetical protein PRO0823

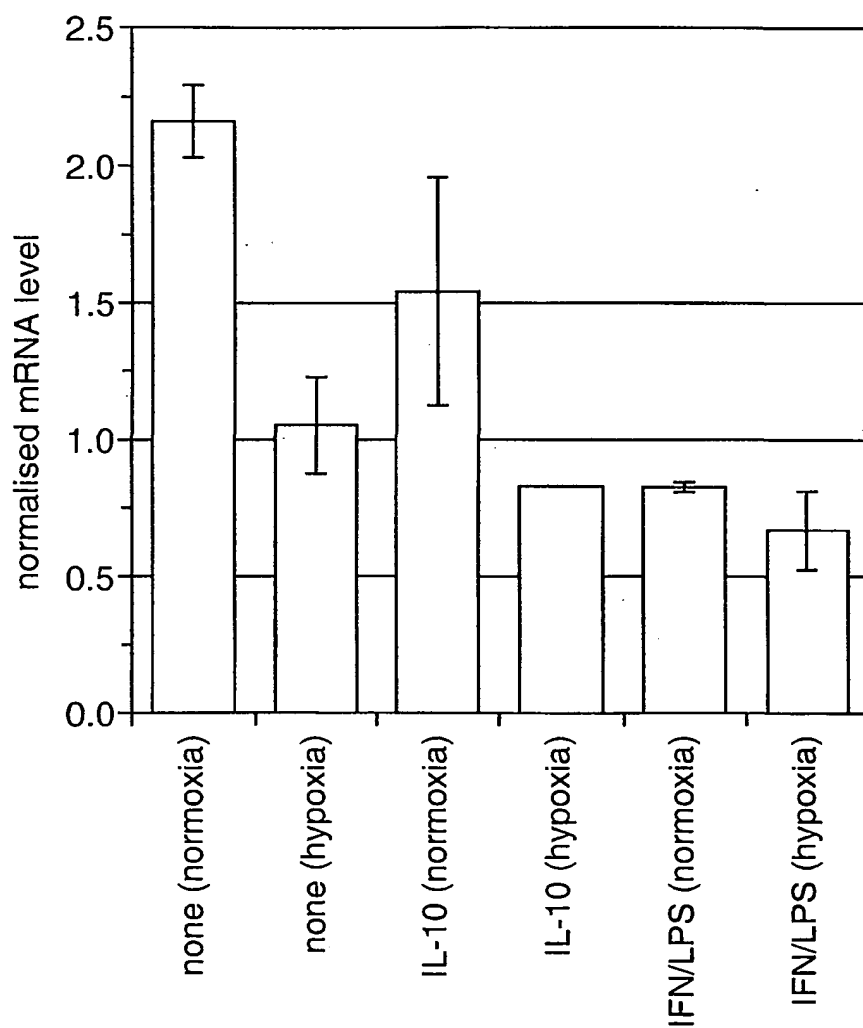


84/96

FIG. 55e

p1F10/ SeqID:6

Hypothetical protein DKFZp434P0116

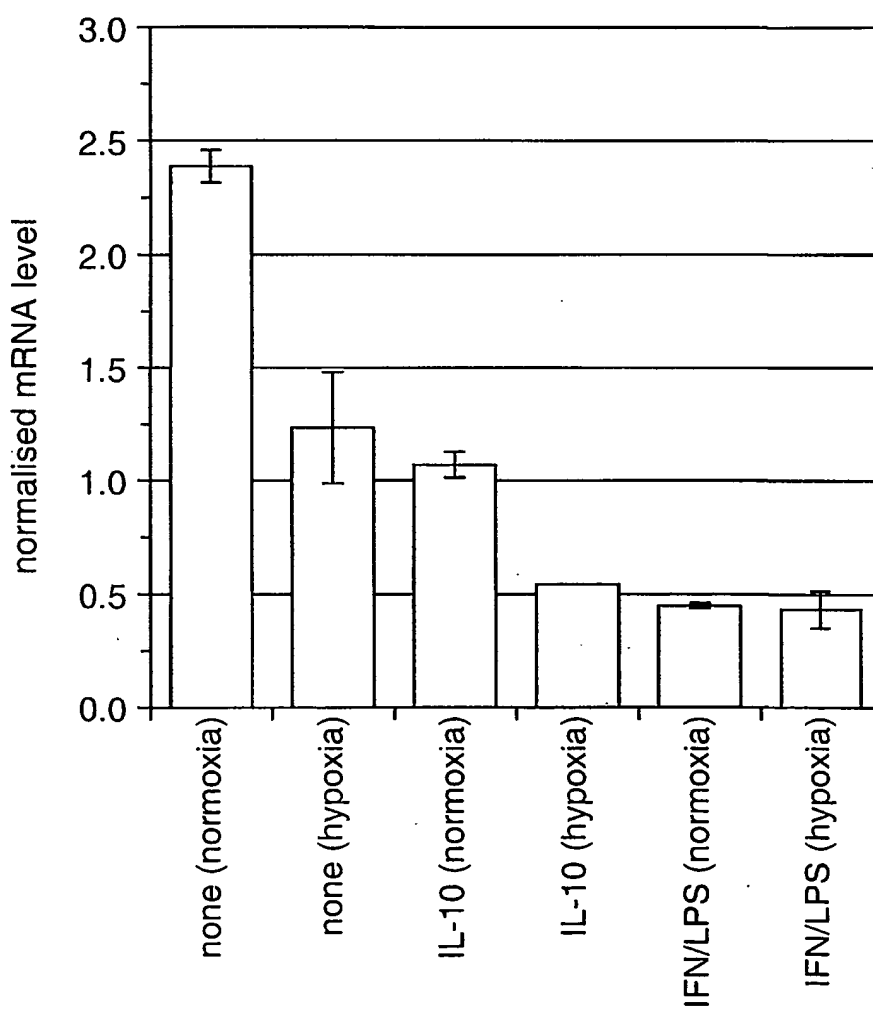


85/96

FIG. 55f

p112/ SeqID:150

cDNA FLJ11302 fis, clone PLACE1009971

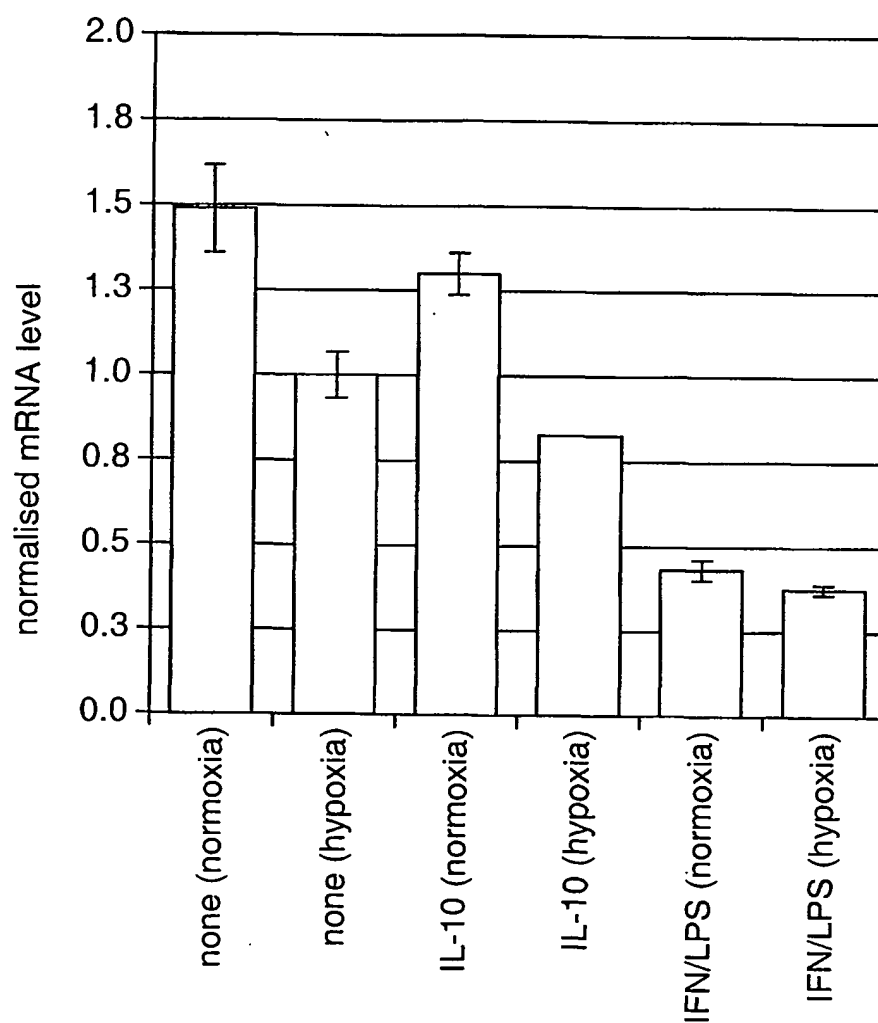


86/96

FIG. 55g

p115/ SeqID:42

Hypothetical protein FLJ10815

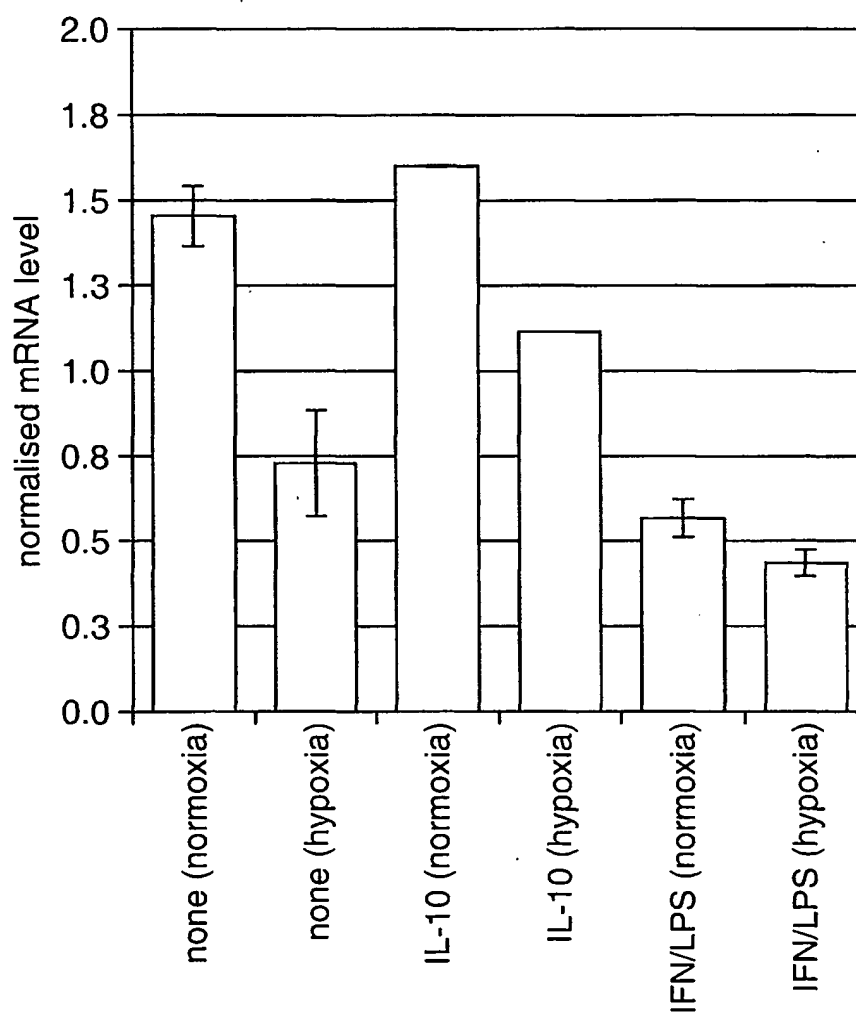


87/96

FIG. 55h

p1G20/ SeqID:204

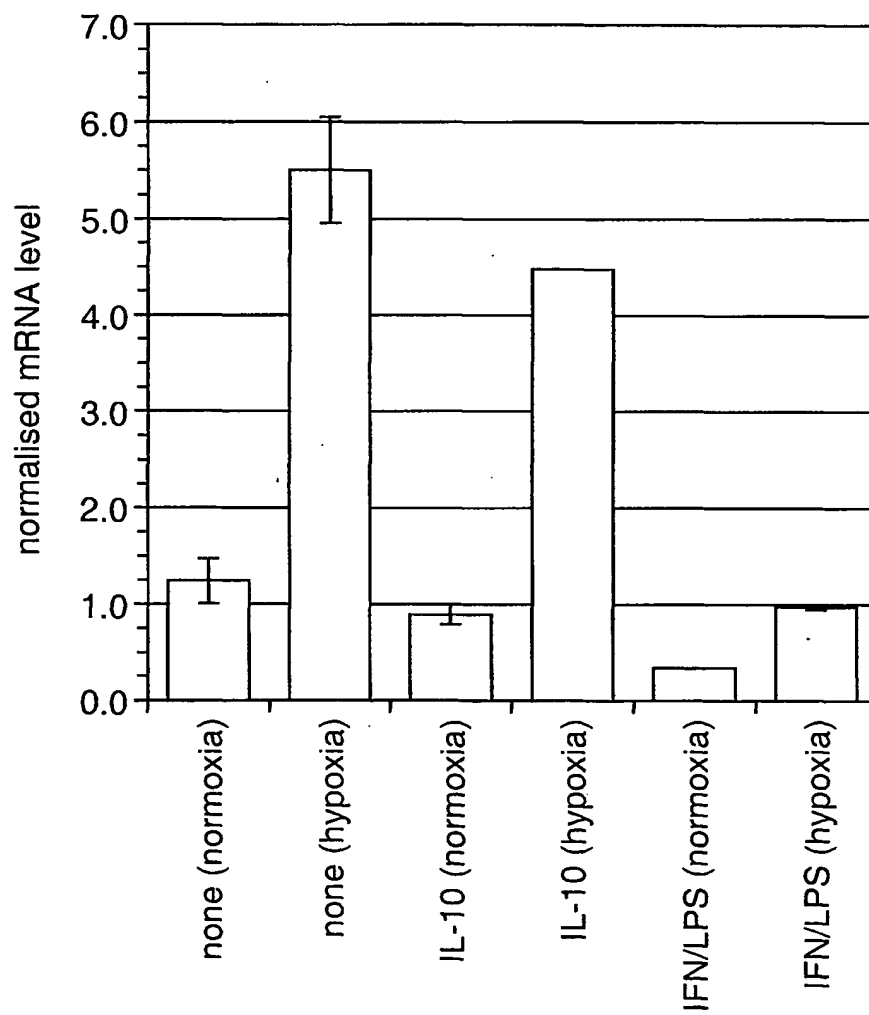
cDNA YO23H03



88/96

FIG. 56a

p1G5/ SeqID:280
MAX-interacting protein 1

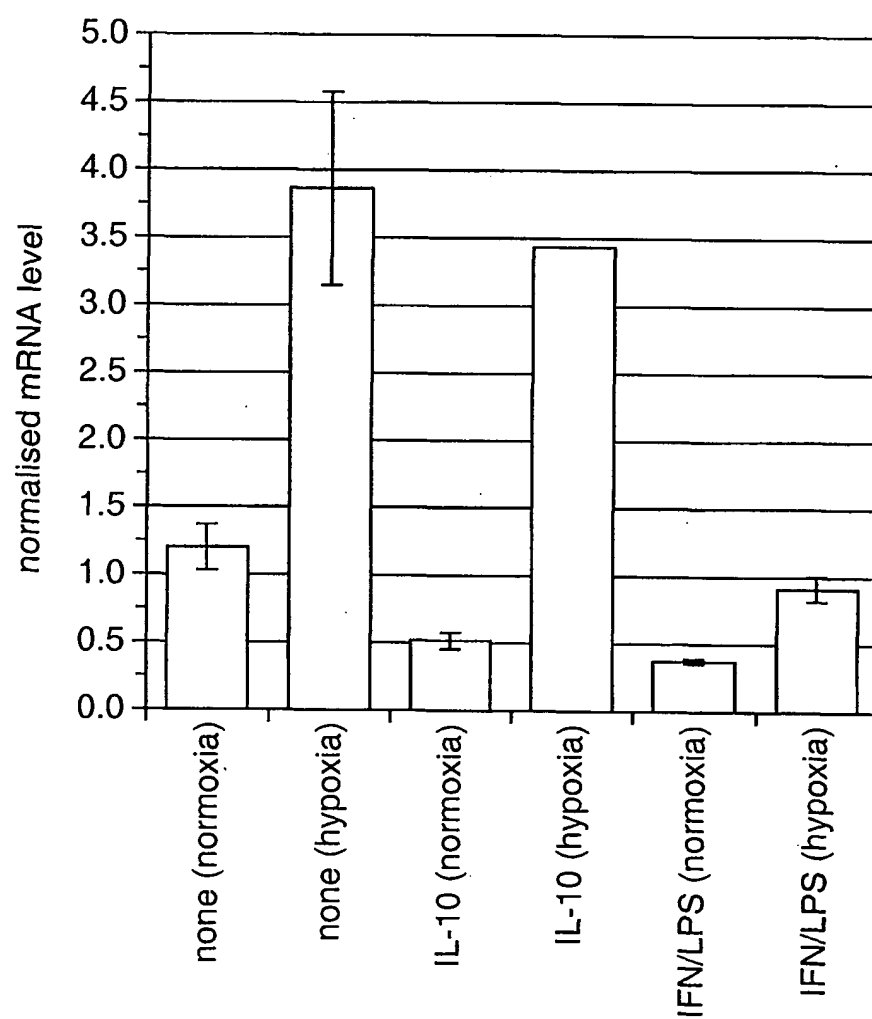


89/96

FIG. 56b

p1D22/ SeqID:120

MAX-interacting protein 1

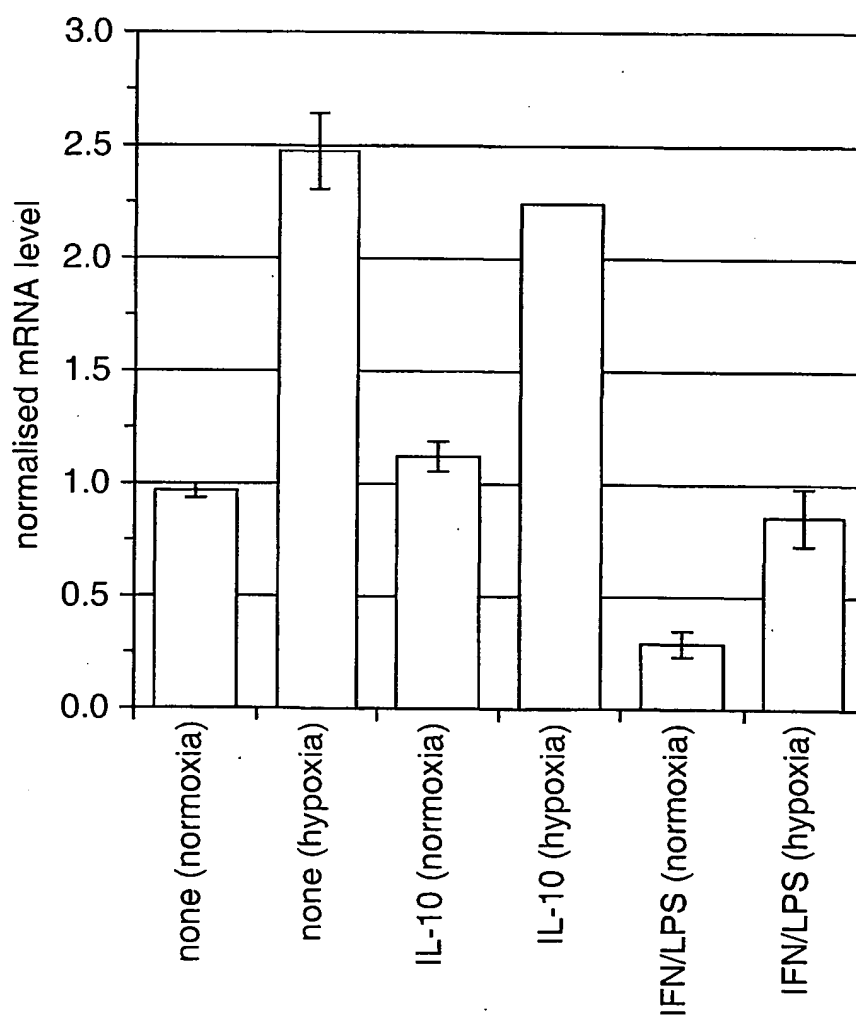


90/96

FIG. 56c

p1G17/ SeqID:316

Early development regulator 2

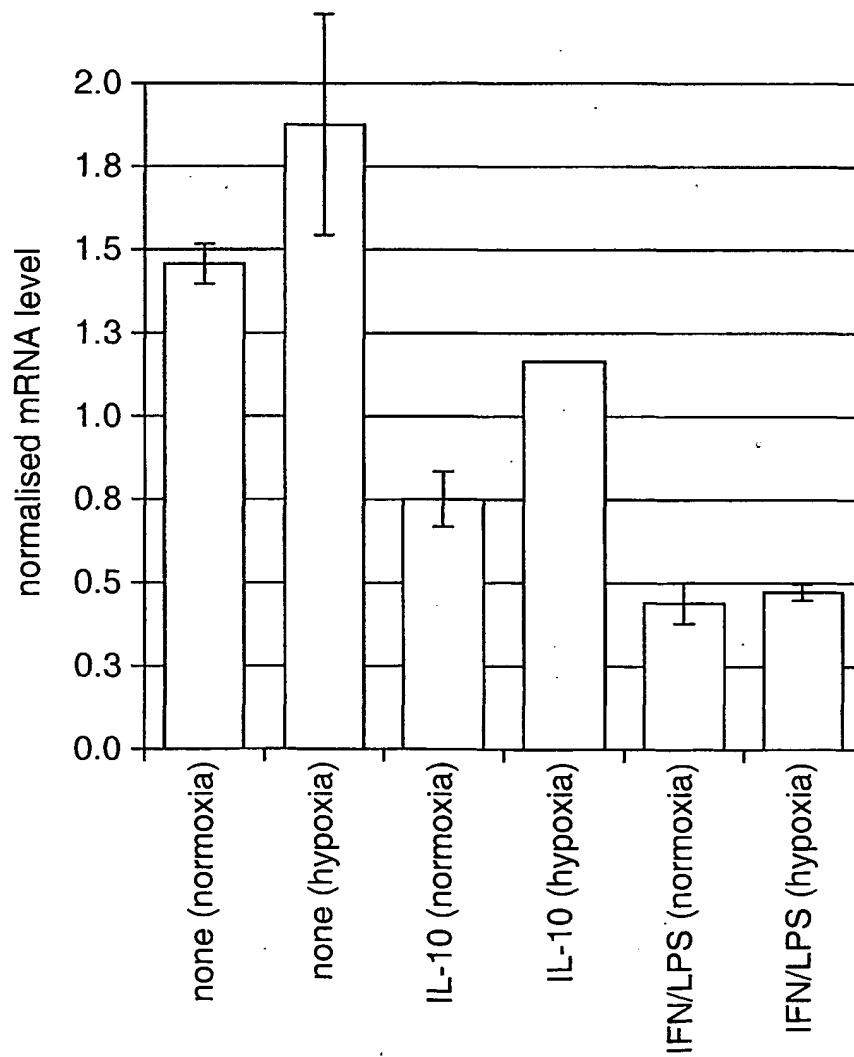


91/96

FIG. 56d

p1G9/ SeqID:306

PI-3-kinase, catalytic, beta polypeptide

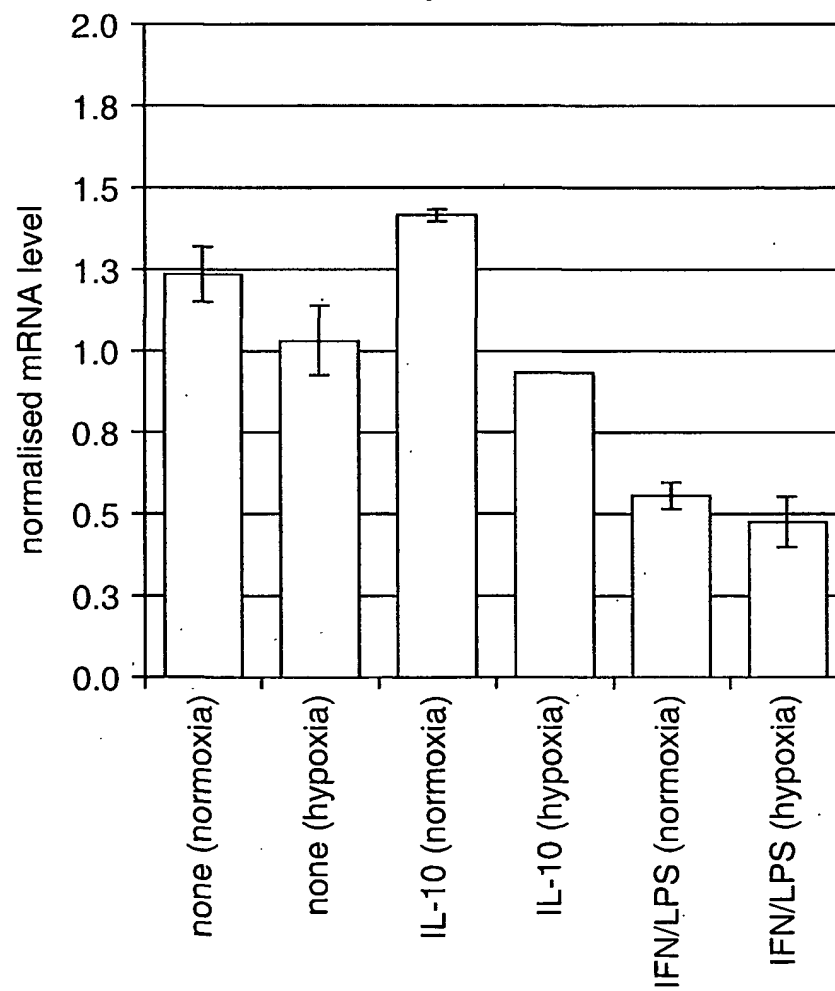


92/96

FIG. 56e

p1K22/ SeqID:420

GPR44

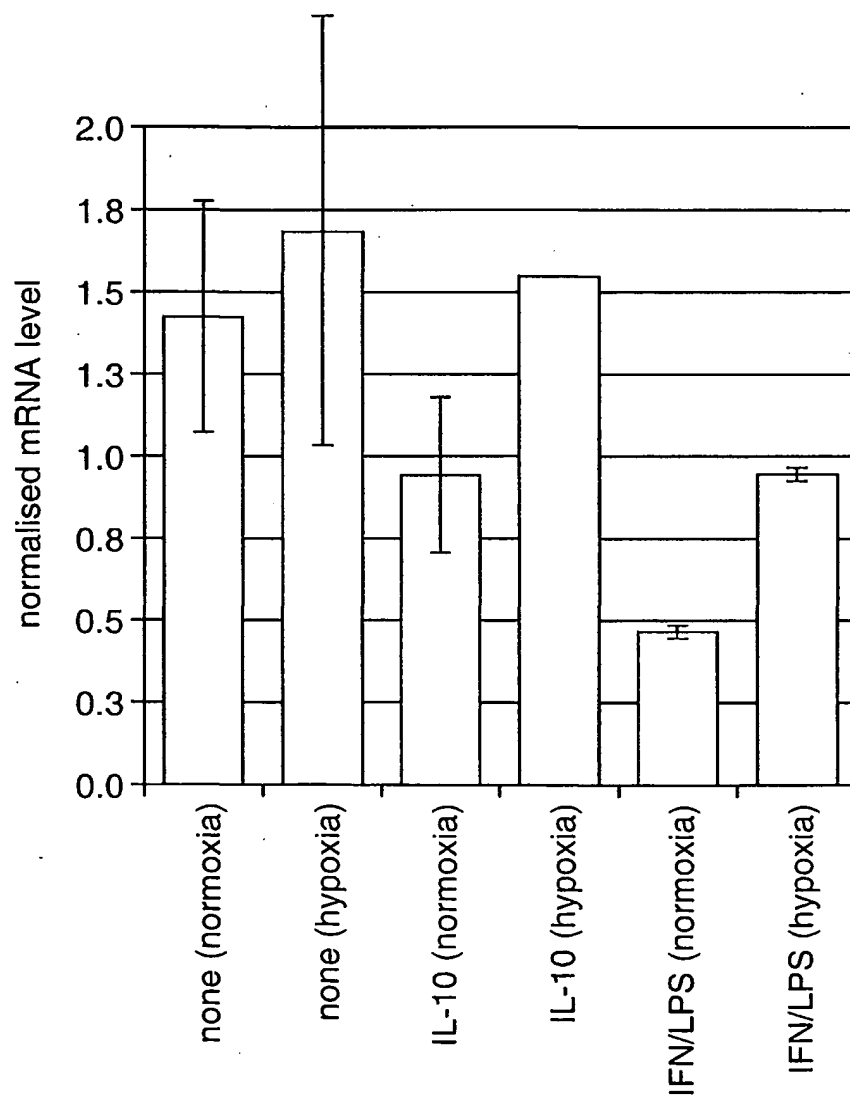


93/96

FIG. 56f

p1C10/ SeqID:376

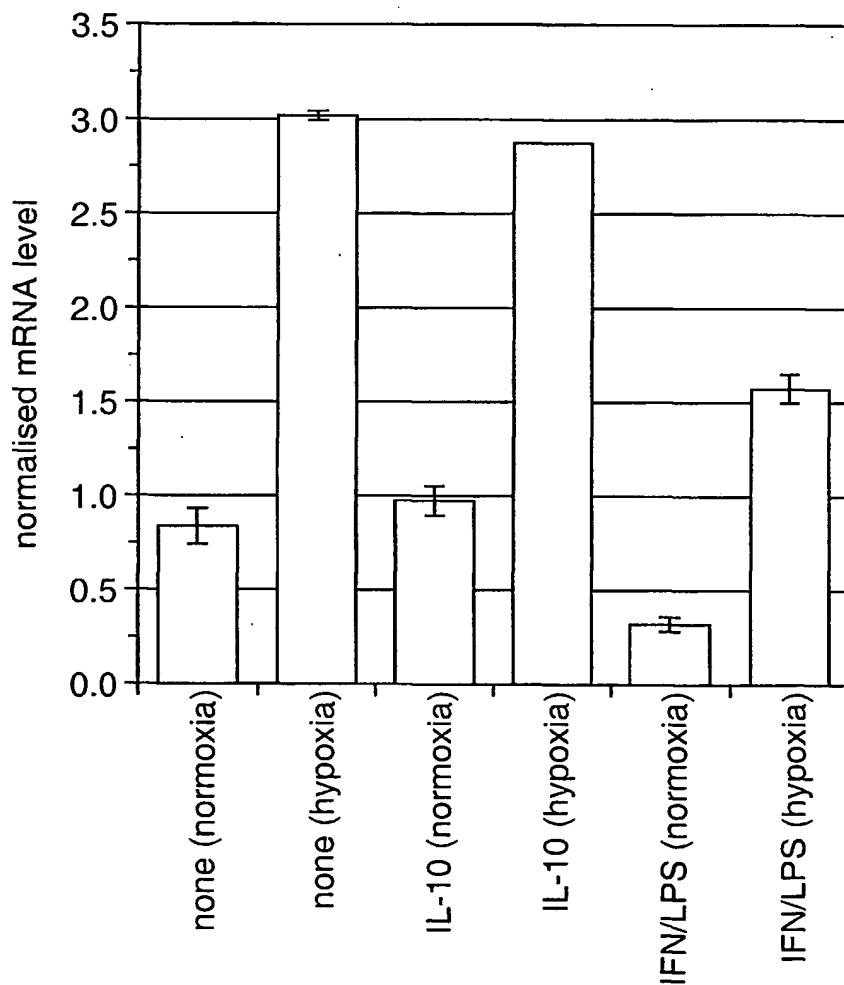
Regulator of G-protein signalling 1



94/96

FIG. 56g

p1D6/ SeqID:68

ERO1 (*S. cerevisiae*)-like

95/96

FIG. 57

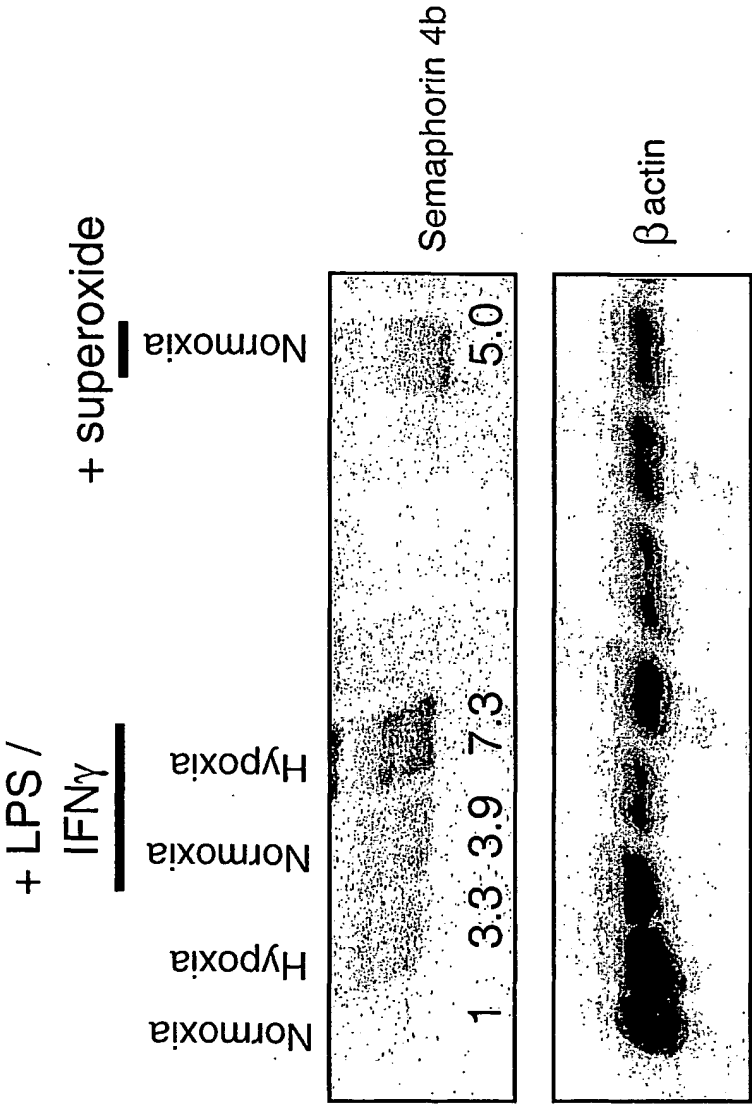


FIG. 58

